Essays in Bayesian implementation

Darina Dintcheva-Bis

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Economics

University of Warwick, Department of Economics

December 2013

Contents

Acknowlegements

I am deeply indebted to my advisor Professor Francesco Squintani for his constant encouragement, guidance and support. I am deeply indebted to my advisor Professor Bhaskar Dutta for his invaluable suggestions. I also want to thank Professor Bhaskar Dutta for his great patience in correcting my errors and for helping me to improve my understanding. None of my advisors are accountable for the contents of this thesis for which I take full responsibility.

Abstract

The purpose of this thesis is to analyze mechanisms that implement a social objective for two environments in which agents have incomplete information about the others' characteristics. Agents' beliefs about the characteristics of any particular agent are common knowledge. We consider the case where monetary transfers or costly signals are undesirable or unavailable. Chapter 1 gives an overview of the contents of the thesis. Chapter 2 studies mechanisms for resolution of bilateral conflict over a prize of common value. This conflict may be settled by a peaceful agreement or may lead to a socially inefficient outcome of war. We model explicitly the cost of war as dependent upon opponents' types which are private information. The social choice function is the probability of peaceful resolution. We assess the chances for peace in the case of no communication and a simultaneous choice by agents whether to agree to a given split proposal. We compare these chances with the probability of peaceful settlement achieved by a mechanism which solicits partial disclosure of private information. We require the truthful revelation of this information to be a dominant strategy for agents in the game induced by the mechanism. In this framework we show that unmediated communication always improves the probability of peace upon the agreement game. In chapter 3 we study a cardinal mechanism for allocation of heterogeneous indivisible goods among agents with private valuations. We assume that agents and the mechanism designer hold the same beliefs about the ex ante distribution of the multidimensional types. We relax the dominant strategy requirement for the truthful revelation to the requirement of Bayesian incentive compatibility. We provide a necessary condition for Bayesian incentive compatibility of such mechanisms for any finite number of goods and agents. We characterize the set of Bayesian incentive compatible mechanisms for the case of three objects and three agents and we analyze efficiency and fairness properties of these mechanisms. In particular, we show that in this framework an ex post efficient and envy-free mechanism may not exist for some systems of beliefs.

Chapter 1

Introduction

Much decentralized decision making involves inference under uncertainty in environments in which agents have preferences over social choices and incomplete information about the preference profile and other agents' information. Implementation theory describes in a formal way the interaction between individuals under the specific rules of an institution or a mechanism. The notion of full implementation requires the exact coincidence of the set of outcomes, prescribed by given solution concept, with the social choice set. A social choice set is partially implemented by the mechanism as long as there exists one equilibrium outcome of the induced game that is socially optimal. In both theoretical and applied works a partial implementation is widely used under different choice of solution concepts (see e.g. [18] by Maskin (2002)).

Palfrey (1993) describes in [24] the Bayesian implementation as a branch of implementation theory where "... individual preferences and information are modelled explicitly, beliefs are updated according to Bayes' rule, outcomes are evaluated according expected utility theory..." and all agents choose actions that maximize expected utility, given strategies of the other agents. Following the work [20] of Myerson (1985), agents' private information is summarized by the notion of a type, where each agent knows his own type but may be unsure about the types of other agents. Beliefs of agents are encoded by probability distributions of type profiles. Equilibrium concept for the game with incomplete information induced by the mechanism is Harsanyi's (1967:1968) Bayesian equilibrium. A social choice function is Bayesian incentive compatible if and only if truthful revelation of private information can arise as a Bayesian equilibrium of some mechanism (see [25] by Palfrey (2002)). A social function is Bayesian implementable if it is Bayesian incentive compatible with respect to all beliefs of all agents. A revelation mechanism may have undesirable equilibria in addition to the truthful one.

We characterize Bayesian mechanisms that implement social choice sets for two different environments with private values and independent types, in order to explain aspects of mechanism selection.

In the first environment we evaluate from the ex ante perspective, i.e., before the parties know their types, alternative mechanisms aiming to resolve a conflict. These mechanisms partially implement the social choice in a Bayesian Nash equilibrium of the induced game.

Application of mechanism design to resolution of conflicts was pioneered by Banks (1990), Fearon (1995), Warneryd (2003), Bester and Warneryd (2006), and Fey and Ramsay (2009). These authors study the design of political and economic institutions to prevent or resolve any type of destructive conflict. A number of processes can be used to resolve a conflict, dispute or a claim. Dispute resolution processes are alternatives to having a war resolve the conflict or handling the dispute in the court system. These processes can be used to resolve conflicts in such areas as interstate relations, family matters, or workplace and contracting practices. Dispute resolution processes are cheaper and usually generate a solution faster than war resolution or litigation. The most common forms of dispute resolution processes are negotiation, mediation, and arbitration. The research on conflict resolution processes focus on conditions under which negotiation or mediation are chosen and the general effectiveness of these processes.

A part of negotiation is bargaining which has been modelled and extensively studied in the economic literature in both cooperative and non-cooperative setups (extensive surveys are provided e.g. by Muthoo (1999) and by Napel (2002)). The equilibrium settlement is well understood in bargaining situations where parties' disagreement payoffs are common knowledge. In his seminal work on axiomatic bargaining theory, Nash (1950) provides a solution of the two-person bargaining problem where the disagreement payoffs are known to both players. The literature on bilateral bargaining in the environment with two-sided uncertainty about outside options is scarce (see e.g. [26] by Sanchez-Pages (2012)). An usual assumption in this literature is that a negotiator faces more than one candidate to reach an agreement with and an outside option for the negotiator become these alternative opponents and prizes. This is an appropriate model in the context of bilateral trade but certainly not a realistic model for all dispute resolution processes. The study of bargaining with incomplete information can be conducted in the framework of Bayesian mechanism design rather than modelled by a sequence of offers and counteroffers. The mechanism design provides a tool for characterization of the set of attainable outcomes for any particular environment and determines the optimal method for conflict resolution from the point of view of the designer.

Goltsman, Horner, Pavlov and Squintani (2009) study in [14] all three different classes of communication procedures: arbitration, mediation and negotiation, in the context of the Crawford and Sobel (1982) model of cheap talk with two players, the informed party and the decision-maker. In this setting, authors find that "mediation performs better than negotiation when the conflict of interest is intermediate, whereas a mediator is unnecessary and two rounds of communication suffice when the conflict of interest is low". These findings about the potential and the limitations of mediation and negotiations are confirmed by some empirical studies of international negotiations and mediations (see e.g. [5] by Bercovitch and Jackson (2001)). Goltsman et al. pose the question about the relative efficiency of negotiations and mediation beyond the classical framework of Crawford and Sobel's model.

In chapter 2 we study the efficiency of alternative forms of resolving disputes between two parties that are uncertain about the cost of war and win probability. We measure efficiency of peace promoting mechanisms by the *ex ante* probability of peaceful agreement. The designer's choice in this environment is a settlement that maximises of the *ex ante* probability of peace. In our model of bilateral conflict the uncertainty about the cost of war and win probability is two-sided, and the cost of war is endogenously determined by the types of opponents. We consider continuous types representing the war capability of contestants. We estimate the probability of peace for an agreement game without communication and for unmediated communication with partial revelation of private information. We show that a peace talk game improves upon the game without communication for all parameters of the model. We study peace benefits over direct communication that a strategic mediator may achieve. We show that some mediation programs with partial revelation of private information do not improve upon unmediated communication with the same message space. We find that these mediation programmes do not introduce sufficient noise into the communication between parties and fail to prevent "leakages" of information to the opponents.

Chapter 2 is organized as follows. In the first section we present a model of conflict in which the loss of welfare in the case of war is proportional to the types of opponents. In section 2 we estimate the probability of peace in the case of lack of communication between players and exogenous split proposal. Section 3 studies the probability of peace in the case of unmediated communication. Section 4 studies the probability of peace for a class of mediation programs with partial revelation of private information. The final section discusses some limitations of our approach and relations to the literature on conflict resolution.

The goal of the mechanisms we consider in chapter 3 is to allocate heterogeneous indivisible goods to agents with private values and common prior belief about preferences of others.

The first version of the problem of allocation of indivisible goods to self-interested agents with privately known preferences is introduced by Hylland and Zeckhauser (1979). In the paper "The Efficient Allocation of Individuals to Positions" Hylland and Zeckhauser propose a procedure for allocation of individuals to positions with capacity constraints. The preferences of individuals are unknown and there are no monetary compensations in this model. It is assumed that each individual truthfully reports to the mechanism the utility level that he receives from each position. Lotteries of probability shares are assigned to individuals using a pseudomarket for these shares. Individual's budget constraint is given by the condition that the sum of his probability shares must equal one. The expected assignment of lotteries is ex ante efficient and envy-free. However, this procedure is not incentive compatible. It presumes that individuals can not affect the competitive equilibrium price of probability shares. Next papers on probabilistic mechanisms adopt an ordinal approach and elicit agents' ordinal preferences.

The standard operator used to extend preference relations over objects to preferences over lotteries is the first order stochastic dominance (sd) introduced in this context by Bogomolnaia and Moulin (2001) in [7]. The content of definition of properties of assignment rules like efficiency, no-envy and strategy proofness are affected by this choice of lottery dominance criterion. Two main probabilistic assignment rules have been considered in the literature. The Random Serial Dictatorship mechanism is strategyproof and anonymous, but only ex-post efficient. It is widely used in applications and studied among others in [1] and [2] by Abdulkadiroglu and Sonmez (1998: 1999). However, Bogomolnaia and Moulin provide in [7] an example where the random serial dictatorship assignment is first order stochastically dominated for all agents by another random assignment. An expected assignment is ordinally efficient or "sd efficient" if no other assignment stochastically dominates the given one for all agents. An expected assignment is "sd envy-free" if according to each agent's preferences his assignment sd dominates the assignment of any other agent. More demanding ordinal efficiency is achieved by the Probabilistic Serial mechanism (PS) introduced by Bogomolnaia and Moulin (2001). It is the only ordinal mechanism which is sd efficient and sd envy free. Efficiency gain of the PS mechanism comes with incentives for agents to misrepresent their ordinal preferences. Sd strategy-proofness of an ordinal rule is the property that a truthful report of preferences guarantees at least as desirable lottery in the stochastic dominance sense as any lottery received by misreporting. Bogomolnaia and Moulin (2001) show that the probabilistic serial mechanism is not sd strategy-proof. Objectives that are achieved by the assignment rule with respect to potentially false preferences are irrelevant to the properties, as sd efficiency and sd no-envy, with respect to the true preferences. This impedes the implementation of the probabilistic serial rule in applications.

Katta and Sethuraman (2006) introduce an extension to the original PS mechanism by allowing agents to be indifferent between objects. The authors show in [8] that when agents are allowed to report indifference it is impossible for even a weak strategy-proof mechanism to find an expected assignment that is both ordinally efficient and envy-free. The concerns for the incentive properties of the PS mechanism may be severe for implementation in small assignment problems. Kojima and Manea show in [17] that truthful reporting of ordinal preferences is a weakly dominant strategy in problems where there are sufficiently many copies of each object.

The question of existence of efficient Bayesian incentive compatible mechanisms in the environment with private valuations and quasi-linear utility has attracted a lot of attention in the literature. In the case where monetary transfers or costly signals are possible, the literature provides characterisation of interim efficient and Bayesian incentive compatible mechanisms. Mechanisms with these properties have been characterised in the context of bilateral trade, auctions, and public good problems. Ohseto (2006) characterises in [23] the set of strategy-proof and envy-free mechanisms of allocating indivisible goods when monetary compensations are possible.

One direction of research on the optimal allocation mechanisms without monetary compensations is represented in [8] by Chakravarty and Kaplan (2013) and in the recent paper [4] by Ben-Porath, Dekel, and Lipman (2013). Chakravarty and Kaplan allow the designer to receive costly and socially wasteful signals from agents. Authors find conditions where allocating the goods randomly is optimal and study cases where mechanisms as a contest or a contest with a bid cap are superior. Ben-Porath, Dekel, and Lipman assume that the mechanism designer can verify the type of agents at a cost and characterise the class of optimal Bayesian incentive compatible mechanisms, showing that the mechanism with favoured agent is optimal.

In comparison, there are a handful of papers on allocation mechanisms in environments without monetary transfers or costly signals. Tayfun Sonmez and M. Utku Unver (2010) consider in [14] a mechanism in which each agent maximises his expected utility given a system of beliefs. These beliefs are given through a joint probability distribution function denoting the probability that the lowest bid required to receive an object (the market-clearing price) will be less than or equal to p. Similar to the Hylland and Zeckhauser's procedure in [15], it is assumed that agents are price takers given this system of beliefs. Miralles (2012) studies in [19] Bayesian incentive compatible cardinal mechanisms for efficient allocation of two ex ante identical objects between two symmetric agents with independent private valuations. The author assumes that each valuation vector is drawn from the same distribution function and does not constrain the number of objects allocated per agent. It is shown that the mechanism which maximises the unweighted sum of agents' *ex ante* expected payoffs is a combination of lotteries, auctions and insurance.

We consider the allocation problem in the environment where monetary compensations or costly signals are undesirable or unavailable. Examples of these class of problems are the assignment of offices among employees and children's placement in public schools. We assume that agents and the mechanism designer share a common belief about the ex ante distribution of preferences which are private information. We study allocation mechanisms that treat both agents and objects symmetrically. No pre-existing ordering is assumed neither for the set of agents nor for the set of objects. A variety of formal criteria of economic justice have been studied in the social choice literature (see [15] by Thompson (2007) for an overview). The "no envy" criterion is proposed by Foley (1967) in [12]. An allocation is envy free if each agent likes his own assignment at least as much as the assignment of any other agent. We consider the existence of a mechanism that allows the social designer to obtain an envy free allocation of objects and simultaneously achieves the social goal of efficiency. In this environment we characterize the set of Bayesian incentive compatible assignment mechanisms.

Chapter 3 is organized as follows. In section 1 we explain the motivation for the study of allocation problem in the environment where monetary compensations are not available and agents hold beliefs about the distribution of preferences. In section 2 we describe an assignment mechanism that elicits cardinal information about agents' preferences and assigns a probability distribution over the set of feasible allocations for each preference profile. In section 2 we provide necessary conditions for incentive compatibility of the mechanisms for any finite number of agents. In section 3 we characterize the set of incentive compatible mechanisms for the case of three agents. In section 4 we provide some necessary conditions for regularity of the optimization problem related to the interim efficient mechanisms. In section 5 we provide examples of the set of ex post efficient and envy-free assignments for the mechanism. In the last section we show that the ex post envy-freeness property may be incompatible with ex post incentive efficiency of the mechanism. We provide an example of common belief where the set of ex post incentive efficient and ex post envy-free assignments for the mechanism is empty.

References

[1] A. Abdulkadiroglu, T. Sonmez *"Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems,* Econometrica 66(1998): 689-701.

[2] A. Abdulkadiroglu, T. Sonmez *House Allocation with Existing Tenants,* Journal of Economic Theory, October 1999,88(2): 233-260.

[3] J.S. Banks *Equilibrium behavior in crisis bargaining games,* American Journal of Political Science, 34 (1990): 599-614.

[4] E. Ben-Porath, E. Dekel, and B.L. Lipman. *Optimal Allocation with Costly Verification,* http://people.bu.edu/blipman/ (2013).

[5] J. Bercovich, R. Jackson *Negotiation or Mediation? An Exploration of Factors Affecting the Choice of Conflict Management in International Conflict,* Negotiation Journal, 17 (2001): 59-77.

[6] H. Bester, K. Warneryd *Conflict and the Social Contract,* Scandinavian Journal of Economics, 108 (2006): 231-49.

[7] A. Bogomolnaia, H. Moulin. *A New Solution to the Random Assignment Problem,* Journal of Economic Theory 100 (2001): 295-328.

[8] S. Chakravarty, T.R. Kaplan *Optimal allocation without transfer payments,* Games and Economic Behavior 77 (2013): 1-20.

[9] J. Fearon *Rationalist Explanations for War,* International Organization, 49 (1995): 379-414.

[10] M. Fey, K. Ramsay, *Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict,* American Journal of Political Science, 55(1) (2011): 149-169.

[11] M. Fey, K. Ramsay, *Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types,* Review of Economic Design, 13(3) (2009): 233-250.

[12] D. Foley *Resource allocation and the public sector,* Yale Economics Essays, 7 (1967): 45-98.

[13] M. Garfinkel, S. Skaperdas, *Economics of Conflict: An Overview,* in Todd Sandler and Keith Hartley, editors, Handbook of Defense Economics, Vol. II (2007) Elsevier.

[14] M. Goltsman, J. Horner, G. Pavlov, F. Squintani *Mediation, arbitration and negotiation,* Journal of Economic Theory, 144(4) (2009): 1397-1420.

[15] A. Hylland, R. Zeckhauser, *The Efficient Allocation of Individuals to Positions,* Journal of Political Economy 87 (1979): 293-314.

[16] A. Katta, J. Sethuraman *A solution to the random assignment problem on the full preference domain* Journal of Economic Theory 131 (2006): 231-250.

[17] F. Kojima, M. Manea *Incentives in the probabilistic serial mechanism* Journal of Economic Theory 145 (2010): 106-123.

[18] E. Maskin,T. Sjostrom, *Implementation Theory,* in K. Arrow, A. Sen, and K. Suzumura, (eds.), Handbook of Social Choice and Welfare, Vol. I, Elsevier (2002): 237-288.

[19] A. Miralles, *Cardinal bayesian allocation mechanisms without transfers,* Journal of Economic Theory 147 (2012): 179–206.

[20] R. B. Myerson *Bayesian equilibrium and incentive-compatibility: An introduction,* Social goals and social organization: Essays in memory of Elisha Pazner (1985): 229-260.
$https://129.105.97.149/research/math/papers/548.pdf$

[21] A. Muthoo *Bargaining theory with applications,* Cambridge University Press, 1999.

[22] S. Napel *Bilateral bargaining: theory and applications,* Vol. 518, Springer, 2002.

[23] S. Ohseto, *Characterization of strategy-proof and fair mechanisms for allocating indivisible goods,* Economic Theory 29 (2006): 111-121.

[24] T. Palfrey, S. Srivastava *Bayesian implementation,* Harwood Academic Publishers, Switzerland (1993).

[25] T. Palfrey, *Implementation theory,* In R.J. Aumann, S. Hart, eds. *Handbook of Game Theory*, Volume 3 (2002): 2273-2326.

[26] S. Sanchez-Pages *Bargaining and conflict with incomplete information,* In M.R. Garfinkel and S. Skaperdas, eds. *The Oxford Handbook of the Economics of Peace and Conflict*, Oxford University Press, 2012.

[27] T. Sonmez, M. U. Unver, *Course bidding at Business Schools,* International Economic Review 51(2010): 99-123.

[28] W. Thompson, *Fair Allocation Rules* Chapter 21 in Handbook of Social Choice and Welfare, vol. 2 (2011): 393-506.

[29] K. Warneryd, *Information in Conflicts,* Journal of Economic Theory 110 (2003): 121-136.

Chapter 2

Peace talks with interdependent cost of conflict

The class of models that is central to current thinking about an international dispute between states is games with two-sided incomplete information. At any stage of negotiations each party can withdraw unilaterally and take up its war payoff. The theoretical literature on international conflict reveals the role of private beliefs of involved parties about the war payoff structure in the shaping of any settlement available through some peaceful process. The literature points out (see [8] by Garfinkel and Skaperdas (2007)) that the outcome of a military conflict between two parties is subject to much uncertainty. The general assumption is that the only source of vacillation are the probabilities of winning a war for each player, while in practice there are multiple sources of uncertainty. One of them is the value of the prize for the prospective winner. The possibility of peaceful resolution of conflict in the case of independent valuations and uncertainty about the opponent's fixed cost of war is studied by Fey and Ramsey (2011) in [5].

In this paper we incorporate this source of cost uncertainty in a situation with interdependent valuations. We model explicitly the cost of a conflict as proportional to the fighting capacities of opponents. These capacities are part of the private information that directly affects payoffs of both disputants in the case of conflict.

Therefore, the outcome of the bilateral bargaining in our model is determined by an outside option for which parties do not have complete information.

We study the Bayesian-Nash equilibria in the *crisis bargaining game* for three types of settlement procedures: a split proposal game, an unmediated communication between parties, and strategic mediation by a third party.

2.1 The model

2.1.1 Structure of private information

Two players contest a prize of common value one. The dispute may lead to a war or litigation. Only a peaceful settlement is *ex post* efficient. The winner of the outright conflict takes a prize of value less than one. Players value *ex post* the war prize identically. The probability of winning the war prize is determined by the relative strength of players. For the sake of simplicity, we assume that players are *ex ante* symmetric. The strength t_i of player i in a potential military conflict is a realisation of a random variable in T and is referred to as type of player i. Each player privately observes his own type. The information of a player about the type of his opponent is summarized by the cumulative probability distribution function $F(t)$ on T. Function $F(t)$ is common knowledge. It is assumed that the realisation of one player' s type does not affect the likelihood of the other player' s types. Formally, the joint distribution function satisfies $F(t_i,t_j) = F(t_i)F(t_j)$ for any $(t_i,t_j) \in T \times T$ where $F(t_i)$ and $F(t_j)$ are marginal distribution functions. We assume thereafter that types of both players are independently and identically drawn from the interval $T = [0,1]$ with uniform probability distribution. A uniform distribution of types is consistent with the assumption of lack of knowledge about opponent's type. We assume that the probability distribution of types after restricting the support is again uniform.

2.1.2 The outside option

At the interim stage an alternative of a peaceful agreement is only unilateral or bilateral initiation of an outright confrontation. The outcome of the confrontation is interpreted as probability of winning the whole prize and not as a split of the prize as in Rubinstein's alternating-offer bargaining game. We denote player i's probability of winning the prize by $p(t_i,t_j)$ and player j's probability of winning by $1-p(t_i,t_j)$. We consider the probability that contestant i gets the prize after his effort level has been exerted and the fighting capability t_i has been obtained. The fighting capability of contestant is a noisy predictor of the outcome of the war or, in a legal context, the litigation. The decision whom to give the prize is determined both by the relative strength of players and by the fairness related type of the decision-making institution or process. The prize lottery faced by each contestant is given by probability measure $p(t_i,t_j)$ on $T \times T$:

$$p(t_i,t_j) = \begin{cases} p & \text{for} \quad t_1 > t_2 \\ \frac{1}{2} & \text{for} \quad t_1 = t_2, \\ 1-p & \text{for} \quad t_1 < t_2 \end{cases}$$

where $p > 1-p$, that is, $p > \frac{1}{2}$. Parameter p is interpreted as a degree to which the relative strength of players is a noisy predictor of the award decision. In a legal context, p is interpreted as a degree to which property rights are defined by the relative strength of arguments (see [2] by Bester and Warneryd (2006)). It is assumed that $p < 1$, that is, property rights are not perfectly defined.

We assume that any possible peaceful settlement is efficient. In order to limit the range of potential settlements that are acceptable by the opponents we assume that they are both risk averse. In the model with win probability $p(t_i,t_j)$, where $t_i > t_j$ and p is the payoff of player i, a peaceful split $(p, 1-p)$ would always be accepted if the relative strength of players was a common knowledge. However, each player has a constant, non decreasing in his type, incentive to misrepresent this type. According to Fey and Ramsay (2009) (see [7]), the combination of uncertainty about the other player's strength in war and the incentive to misrepresent private information has been identified in the literature as a central cause of war. Fearon (1995) points out in [4] that the bargaining might not prevent a war if the prize at stake is indivisible. The indivisible probability $p(t_i,t_j)$ of getting the prize can be considered as an additional war favouring condition to the indivisible prize.

2.1.3 War payoffs

Both players can take an unilateral action and induce war. Player's payoff from this activity is not known *ex ante*. The uncertainty about this option is twofold. Apart from the uncertainty about the probability of winning the prize, players are uncertain about the cost of conflict. A conflict shrinks the value of the prize. The standard assumption in the literature is that a conflict destroys a fixed part of the initial value (see [2] by Bester and Warneryd (2006) and [14] by Warneryd (2010)). This corresponds to the assumption that the war budget is fixed while in practice it is difficult to translate into financial terms and changes with the opportunity cost of doing something else. Hence, it is reasonable to assume that the loss, due to the overall resources expended, is a function of strengths of players involved in the conflict.

In both military and legal context, it is more costly to pick the winner when players differ less in their strength. Hence, the lower is the difference, the more destructive is conflict. Additionally, the value of the prize $\theta(t_i,t_j)$ allocated to the winner is decreasing in both players' strength. We assume that the value of the prize $\theta(t_i,t_j)$ allocated to the winner satisfies conditions

$$\theta(t_i,t_j) > \theta(t_i',t_j) \quad if \quad t_i' > t_i, \quad and \quad \theta(t_i,t_j) > \theta(t_i,t_j') \quad if \quad t_j' > t_j.$$

It is easy to check that when the value of the prize depends solely on the difference $t_i - t_j$ in fighting capacities of opponents then for any feasible values of parameters the expected war payoff is sufficiently low for the peaceful split $(1/2, 1/2)$ to be unanimously accepted. That is why we make a more sound assumption that the destruction technology is given by some function

$$\theta(t_i,t_j) = 1 - \alpha t_i\, t_j, \tag{2.1}$$

where $0 < \alpha < 1$. This choice of technology reflects the observation that for any given distribution of fighting resources t_1 and t_2 a war between players of equal strength is the most destructive. This is the case because for any fixed total resource $S = t_1 + t_2$, the constrained maximum $max_{\{t_1+t_2=S\}}\{t_1\ t_2\}$ is achieved when $t_1 = t_2$. Besides, a war between stronger opponents shrinks more the prize value given by (2.1).

2.2 Peaceful settlement without communication

In this paragraph we calculate the *ex ante* probability of peaceful settlement in the Bayesian-Nash equilibria of a split proposal game. This mechanism does not solicit any private information and assumes that players do not communicate. The set of feasible outcomes in the split proposal game represents values of the prize allocated to players.

Definition 1 *The set of feasible outcomes is* $Y = \{(y_1, y_2) : y_1 + y_2 \leq 1\}$. *The set* Y^e *for which* $y_1 + y_2 = 1$ *is called the set of efficient outcomes.*

Players are proposed a split $(x, 1-x) \in Y^e$. At the time of proposal there is no common or private knowledge about the relative strength of players. After observing their own type, players simultaneously choose whether to agree or not to the given split proposal. If both players accept the split then war is prevented, otherwise they fight. We assume that no other peace promoting mechanism is available.

A pure strategy $\sigma_i(x,t)$ of player i in the split proposal game specifies one of the two responses, 'accept' or 'reject', for each split proposal and each potential type. Before responding to the proposal, each player estimates the conditional probability distribution of his war payoff using the prior probability distribution F. Beliefs of players about war payoffs are formed on the base of both parameters p and α, under the lack of knowledge about the relative strength.

The probability that proposal $(p, 1-p)$ or $(1-p, p)$ will be jointly accepted is zero, although with probability $1/2$ the value $p > 1/2$ is proposed to the higher type. As we show in subsequent paragraphs, the reason is that $1-p$ is the minimal expected war payoff. Hence, there are no values of p and α for which proposal $(p, 1-p)$ or $(1-p, p)$ leads to voluntary peaceful settlement with positive probability.

Denote by $\pi(t)$ the expected war payoff for any player of type t.

Definition 2 *The assessment $(x, \sigma_1(x,t_1), \sigma_2(x,t_2), \pi(t_1), \pi(t_2))$ is a pure strategy Bayesian Nash equilibrium of the split proposal game if*

(i) given the split x, for each player i of type t, the response $\sigma_i(x,t)$ maximises the expected value of the allocated prize, given his belief $\pi(t_i)$,

(ii) for each x, each player's belief $\pi(t_i)$ satisfies Bayes' rule, that is,

$$\pi(t_i) = \int_0^1 p(t_i,t_j)\theta(t_i,t_j)\,dF(t_j).$$

Denote by $\mathscr{P}(x, 1-x)$ the probability of peace with peaceful split $(x, 1-x)$.

Proposition 1 *The probability of peace $\mathscr{P}(1/2, 1/2)$ in the Bayesian Nash equilibrium of the split proposal game without communication is* 1 *for* $p(1-\alpha) < 1/2$.

Proof: Consider beliefs of players regarding their relative strength. The split $(\frac{1}{2}, \frac{1}{2})$ will be accepted by a player who beliefs the opponent to be of the same type because $\alpha > 0$ and $\frac{1}{2} \geq \frac{1}{2}(1-\alpha t^2)$ for any $t \in T$. Let player i beliefs to be of the higher type. The war payoff of the higher type is $\pi(t_i) = p(1-\alpha t_i t_j) < p(1-\alpha t_i^2) < p(1-\alpha)$. Therefore, if $p(1-\alpha) < \frac{1}{2}$, then a peaceful split $(\frac{1}{2}, \frac{1}{2})$ will be accepted by the player who beliefs to be of the higher type. Let player i beliefs to be of lower type. The war payoff of the lower type is $\pi(t_i) = (1-p)(1-\alpha t_i t_j) < (1-p)(1-\alpha t_j^2) < p(1-\alpha t_j^2) < p(1-\alpha)$ because $p > \frac{1}{2}$. Therefore, if $p(1-\alpha) < \frac{1}{2}$, then a peaceful split $(\frac{1}{2}, \frac{1}{2})$ will be accepted by both players, irrelevant of their beliefs about the type of the opponent. Hence, peace can be attained with probability 1 when $p(1-\alpha) < \frac{1}{2}$. $\square$

Hereafter we shall assume that

$$p(1-\alpha) > \frac{1}{2}. \tag{2.2}$$

In subsequent paragraphs we assess the *ex ante* probability of peace in the Bayesian Nash equilibrium of the peace proposal game for every given split x. We shall determine the split which maximizes this chance.

We will be using the following lemma.

Lemma 1 *Expected war payoff $\pi(t)$ of type t is strictly increasing in t for parameters satisfying condition* (2.2).

Proof: Conditional probability of being the higher type for player i of type t_i is t_i because $P\{t_j < t_i\} = t_i$. The expectation of truncated type t_j is

$$E\{t_j | a < t_j \leq b\} = \int_a^b \frac{t_j}{b-a} dt_j.$$

The payoff expectation for player 1 in case of war, when his type is known to be t_1, is given by

$$\pi(t_1) = \int_0^1 p(t_1,t_2)\theta(t_1,t_2)\, dF(t_2) = \tag{2.3}$$

$$= \int_0^{t_1} p(1-\alpha t_1 t_2)\, dF(t_2) + \int_{t_1}^1 (1-p)(1-\alpha t_1 t_2)\, dF(t_2) =$$

$$= 1 - p + (-1 - \frac{\alpha}{2} + 2p + \frac{\alpha p}{2})t_1 + (\frac{\alpha}{2} - \alpha p)t_1^3. \tag{2.4}$$

We show in the next paragraph that expression (2.4) is strictly increasing in t_1 for $t_1 \in (0,1)$ and any pair of parameters α and p that satisfies condition (2.2).

We notice that stationary points of $\pi(t_1)$ satisfy condition

$$\frac{\partial \pi(t_1)}{\partial t_1} = -2-\alpha+4p+\alpha p+3\alpha t^2-6\alpha p t^2 = 0. \tag{2.5}$$

The coefficient in front of the cube of t_1 in expression (2.4) is negative because $\frac{\alpha}{2}-\alpha p = \alpha(1/2-p)$ and $p > 1/2$. Hence, $\pi(t)$ is a concave function. Moreover, $\pi(t_1)$ tends to $-\infty$ when t_1 tends to ∞ and vice versa. Then the negative stationary point $-\sqrt{\frac{4p+\alpha p-\alpha-2}{6\alpha p-3\alpha}}$ of $\pi(t_1)$ corresponds to a local minimum and the positive stationary point $\sqrt{\frac{4p+\alpha p-\alpha-2}{6\alpha p-3\alpha}}$ of $\pi(t_1)$ corresponds to a local maximum. Hence, it is sufficient to show that the positive stationary point is higher than 1. We notice that

$$\frac{4p+ap-a-2}{6ap-3a} > 1 \Leftrightarrow a < \frac{4p-2}{5p-2} \Leftrightarrow a < \frac{2p-1}{\frac{5}{2}p-1}.$$

By (2.2) inequality $a < \frac{2p-1}{2p}$ holds. Additionally,

$$\frac{2p-1}{2p} < \frac{2p-1}{\frac{5}{2}p-1}$$

because $\frac{5}{2}p-1 < 2p$ for any $p<1$. Hence,

$$\alpha < \frac{2p-1}{5/2p-1}.$$

Therefore,

$$\frac{4p+\alpha p-\alpha-2}{6\alpha p-3\alpha} > 1$$

for any parameters α and p satisfying condition (2.2). $\square$

Notice that the minimal expected payoff of a player is $\pi(0) = 1-p < \frac{1}{2}$ and the maximal expected payoff is

$$\pi(1) = p(1-\frac{\alpha}{2}) > \frac{1}{2}.$$

In order to convince the player of type t_i to accept a payoff x we set

$$x \geq \pi(t_i).$$

We denote by $\mathscr{P}(x,1-x)$ the probability of peace with peaceful split $(x,1-x)$. This probability is determined by the joint cumulative distribution function of random variables $\pi(t_1)$ and $\pi(t_2)$. These random variables are independent because t_1 and t_2 are independent. Hence, the joint cumulative distribution is a product of marginal distributions. Therefore,

$$\mathscr{P}(x,1-x) = P(x \geq \pi(t_i) \wedge 1-x \geq \pi(t_j)) = P(x \geq \pi(t_i))P(1-x \geq \pi(t_j)).$$

The following claim is proved in the Appendix.

Proposition 2 *The probability of peace with equal split is*

$$\mathscr{P}(1/2,1/2) = -\frac{4}{3} r \sin^2\left(\frac{\pi}{6} - \frac{1}{3}\arccos\left(\frac{3\sqrt{3}q\sqrt{-\frac{1}{r}}}{2r}\right)\right)$$

where r and q are given by

$$r = \frac{-2-\alpha+4p+\alpha p}{\alpha(1-2p)} \quad \text{and} \quad q = \frac{1}{\alpha}. \tag{2.6}$$

This result shows that the probability of peace with equal split is lower that $1/2$ for some parameter values. For example, if $(p,\alpha) = (3/4,1/4)$ then the probability of peace with equal split is $\mathscr{P} \approx 0.310$, if $(p,\alpha) = (5/6,1/8)$ then $\mathscr{P} \approx 0.267$, if $(p,\alpha) = (3/4,1/6)$ then $\mathscr{P} \approx 0.286$.

The expected war payoff $\pi(t)$ is nonlinear in t and it seems difficult to calculate its cumulative distribution function. However, by Lemma (1) function $\pi(t)$ is monotonically increasing for $t \in [0,1]$. Hence, a linear approximation of $\pi(t)$ over the interval $[0,1]$ is appropriate and has the advantage of being uniformly distributed. We would like to find an upper bound for the probability $\mathscr{P}(x,1-x)$. By using a linear approximation which provides a lower bound for $\pi(t)$, we prove the following result.

Proposition 3 *The probability of peace $\mathscr{P}(x,1-x)$ in the Bayesian Nash equilibrium of the split proposal game without communication satisfies inequality*

$$\mathscr{P}(x,1-x) \leq \frac{(2p-1)^2}{(4p-2-\alpha p)^2} \quad for \quad p(1-\alpha) > 1/2.$$

Proof: Function $\pi(t)$ is concave for $t \in [0,1]$ (see proof of Lemma (1)). Hence, the linear function $f(t) = At+B$ which satisfies conditions $f(1) = \pi(1) = p(1-\frac{\alpha}{2})$ and $f(0) = \pi(0) = 1-p$ provides a lower bound for $\pi(t)$. Then $B = 1-p$ and $A = p(1-\frac{\alpha}{2}) - (1-p) = 2p-1-\frac{\alpha p}{2}$.

Hence,

$$f(t) = At + B = (2p - 1 - \frac{\alpha p}{2})t + 1 - p.$$

The distribution of $f(t)$ is defined by the cumulative distribution function

$$F_f(x) = P(f(t) \le x) = P\left(t \le \frac{x-B}{A}\right) = \tag{2.7}$$

$$= P\left(t \le \frac{x-(1-p)}{2p-1-\frac{\alpha p}{2}}\right) = F_t\left(\frac{x-(1-p)}{2p-1-\frac{\alpha p}{2}}\right).$$

For example, an upper bound for the probability that a payoff of $1/2$ will be accepted in the equilibrium is given by inequality

$$P(1/2 \ge \pi(t)) \le P(1/2 \ge f(t))$$

where

$$P(f(t) \le 1/2) = P\left(t \le \frac{\frac{1}{2}-(1-p)}{2p-1-\frac{\alpha p}{2}}\right) = F_t\left(\frac{p-1/2}{2p-1-\frac{\alpha p}{2}}\right) = \frac{2p-1}{4p-2-\alpha p}.$$

Then the upper bound for the probability of peace with equal split is

$$\mathscr{P}(1/2, 1/2) \le \left(\frac{2p-1}{4p-2-\alpha p}\right)^2 = \hat{\mathscr{P}}(1/2, 1/2).$$

The accuracy of this estimation is exemplified by the upper bounds $\hat{\mathscr{P}}$ for $(p,\alpha) = (3/4, 1/4)$, $(p,\alpha) = (5/6, 1/8)$, and $(p,\alpha) = (3/4, 1/6)$ which are approximately 0.379, 0.294, and 0.327.

Further, we can find the upper bound for the probability of peace with split $(y, 1-y)$ for any $y \in [0,1]$. By substituting $x = 1-y$ and $x = y$ in (2.7) we obtain

$$F_f(1-y) = F_t\left(\frac{(1-y)-(1-p)}{2p-1-\frac{\alpha p}{2}}\right) = F_t\left(\frac{p-y}{2p-1-\frac{\alpha p}{2}}\right) = \frac{p-y}{2p-1-\frac{\alpha p}{2}}$$

and

$$F_f(y) = F_t\left(\frac{y-(1-p)}{2p-1-\frac{\alpha p}{2}}\right) = \frac{y-1+p}{2p-1-\frac{\alpha p}{2}}.$$

Then

$$\mathscr{P}(y, 1-y) \le \frac{(y-1+p)(p-y)}{(2p-1-\frac{\alpha p}{2})^2} = G(y).$$

Function $G(y)$ has a local maximum at $y = 1/2$ because $G'(y) = \frac{1-2y}{(2p-1-\frac{\alpha p}{2})^2}$ and $G''(y) = -2 < 0$. It is easy to check that this local maximum is also a global maximum for $y \in [0,1]$. Hence, for any $y \in [0,1]$ the probability of peace in the equilibrium of the split proposal game satisfies

$$\mathscr{P}(y, 1-y) \leq G(1/2) = \frac{(p-1/2)^2}{(2p-1-\frac{\alpha p}{2})^2} = \frac{(2p-1)^2}{(4p-2-\alpha p)^2} < 1.$$

Then

$$\mathscr{P}(x, 1-x) \leq \frac{(2p-1)^2}{(4p-2-\alpha p)^2} \quad for \quad p(1-\alpha) > 1/2.$$

□

Corollary 1 *In the game without communication where parameter values satisfy inequality* (2.2) *each split offer leads to peace with probability less than* 1/2.

Proof: Let $p(1-\alpha) > 1/2$. By proposition (3) inequality

$$\mathscr{P}(x, 1-x) \leq \frac{(2p-1)^2}{(4p-2-\alpha p)^2}$$

holds. It is easy to check that inequality

$$\mathscr{P} \leq \frac{(2p-1)^2}{(4p-2-\alpha p)^2} < \frac{1}{2} \tag{2.8}$$

holds for $p(1-\alpha) > 1/2$.

□

2.3 Peaceful agreement in the unmediated peace talk game

In this paragraph we calculate the *ex ante* probability of peaceful settlement in the Bayesian-Nash equilibria of unmediated peace talk game. We assume hereafter that the conflict is not sufficiently costly to secure the peace with probability one. Hence, inequality (2.2) holds.

We consider a cheap talk where the message space is the product $\tau = M \times M$. After learning their own type, players publicly and simultaneously send costless messages $m_i \in M$, $i = 1, 2$. Denote the generic message profile by $m = (m_1, m_2)$.

Messages sent in this mechanism are unverifiable, except for the case of war resolution.

For any given message profile m the mechanism may recommend to players either acceptable or unacceptable split. An acceptable split in the peace talk game is an element from the set of efficient outcomes Y^e that might be accepted with strictly positive probability by both players who observe the message profile. The set of acceptable splits is denoted by X, $X \subset Y^e$. The set X can be represented by the share x for player 1 in recommendations. Recall that the *ex ante* expected payoff of type $t_i = 0$ is $1-p$. Hence, by the symmetry of players, X is equivalent to the interval $[1-p,p]$. The probability of recommendation of acceptable split $x \in [1-p,p]$ when the message profile is m is denoted by $q(m)$, $q(m) \in [0,1]$. The probability of recommendation of unacceptable, war inducing split $x \notin X$, is denoted by $1-q(m)$. A set of decisions is the product set $D = [1-p,p] \times [0,1]$ with generic element (x,q).

Definition 3 *A peace talk mechanism is a pair (τ, f) where the decision rule $f : \tau \to D$ maps each pair of messages into a decision in $D = [1-p,p] \times [0,1]$.*

We assume that no other peaceful mechanism is available. Hence, any unilateral rejection of a split x proposed by mechanism (τ, f) leads to war with probability 1. The outcome induced by the mechanism (τ, f) is either the efficient split x or, in the case of unilateral or bilateral rejection of recommendation, an inefficient outcome from the set $Y \setminus Y^e$ as a result of war.

We consider the game G induced by the mechanism (τ, f). In this game player i adopts a strategy $s_i : T \to M$ which maps each type to a message.

Preferences of players in the game G are given by utility functions $u_i : T \times M \to [0,1]$. Utilities from a messages profile $m = (m_1, m_2)$ are

$$u_1(t,m) = q(m)x(m) + (1-q(m))\pi_1(t)$$

and

$$u_2(t,m) = q(m)(1-x(m)) + (1-q(m))\pi_2(t)$$

where $\pi_1(t)$ and $\pi_2(t)$ denote expected war payoffs of players when the type profile is t. Let $b_0(.)$ be the uniform common prior probability distribution over type profiles T. The game G is described by $(\{1,2\}, T, M, b_0, (u_1, u_2))$. Denote by $s(t_1,t_2) = (s_1(t_i), s_2(t_2))$ the pair of messages generated by the strategy profile (s_1, s_2) when realised types are (t_1, t_2).

A pure strategy Bayesian-Nash Equilibrium of G is a system of type-independent believes $b(.) = (b_1(.), b_2(.))$ at every information set and a pair of strategies s that are interim best response to the strategies used by the other player. We compute the expected payoff U_i of player i with respect to the belief b_i about the type of opponent from interim perspective, after player i learns his own type. Given the belief b_i, the expected payoff for player i of strategy profile s is

$$U_i^{b_i}(t_i) = \int_0^1 u(t_1, t_2, s(t_1, t_2))\, db_i(t_j),$$

where $i \neq j$.

A direct mechanisms is a mechanism in which each player i fully or partially identifies himself by sending a message $m_i \in T$. If $M = T$ then a truth telling strategy for agent i is to reveal precisely its type, that is, $s_i(t_i) = t_i$. A direct mechanism is Bayesian incentive compatible if it has a Bayesian equilibrium $(s_1^*(t_i), s_2^*(t_2))$ such that $s_1^*(t_i) = t_i$ and $s_2^*(t_2) = t_2$. We wish to find out if a decision rule f, such that the expected value of q for the rule f is higher than $1/2$, can be implemented as a Bayesian Nash equilibrium of the game induced by some cheap talk mechanism. By the Revelation Principle for Bayesian equilibrium (see [12]), if a mechanism (τ, f) implements decision rule f in Bayesian equilibrium of the induced game, then the direct mechanism implementing f is Bayesian incentive compatible. Hence, we restrict our consideration to direct mechanisms (τ, f). Fix the equilibrium strategy profile s^* for mechanism (τ, f). Following the notation in [7], we denote by $q^*(m)$ the probability that, after observing each other's message, both players accept the recommended split $x(m)$. Then the value of the prize for player 1 obtained from participation in the mechanism (τ, f) is $x^*(t) = x(s^*(t))$.

We evaluate the quality of the mechanism (τ, f) by the probability $q^*(m)$ with which it is likely to avoid confrontation at the equilibrium, given the updated beliefs of players about the type of the opponent. We assume that these beliefs are formed according to Bayesian rule and updated in the light of observed messages and the recommended split. We assume no commitment of players to the mechanism, therefore it must be optimal *ex post* in the equilibrium of G to accept all peaceful splits proposed.

Let in the mechanism (τ, f) both functions $q(m)$ and $x(m)$ are symmetric across messages. This entails that

$$x(m_1, m_2) = x(m_2, m_1) \quad \textit{and} \quad q(m_1, m_2) = q(m_2, m_1). \tag{2.9}$$

Hence,

$$x(y,y) = 1/2 \quad for\ any\ y \in M.$$

In a pooling equilibrium of the game each player sends the same message irrelevant to his type. The outcome in this equilibrium coincides with the equilibrium outcome of the agreement game without communication.

Our objective is to find a direct mechanism that maximizes the *ex ante* expected probability of peace

$$\mathscr{P} = max_{x(m),q(m)} \int_M \int_M q(m_1,m_2)\, dm_2\, dm_1 \tag{2.10}$$

across all mechanisms that satisfies the following two conditions. In the Bayesian-Nash equilibrium of the induced game G each player's strategy should satisfy the *interim* incentive compatibility (IC) and the *ex post* individual rationality (IR) (participation) constraints that will be defined in subsequent paragraphs with respect to the message space M.

2.3.1 Full disclosure peace talk game

In this subsection we show a property of incentive compatible direct mechanism mechanism (τ, f) for the case where each player's message space is $M = T$, i.e., $M = [0,1]$. This construction implicitly assumes that every type is certifiable, that is, point estimates of fighting capacities of opponents can be certified by external experts.

We consider symmetric mechanisms that satisfy conditions (2.9). We would like to find a mechanism with splits $x(m_1,m_2)$ and probabilities $q(m_1,m_2)$ that maximizes the *ex ante* probability of peace

$$max_{x(m_1,m_2),q(m_1,m_2)} \int_0^1 \int_0^1 q(m_1,m_2)\, dm_2\, dm_1 \tag{2.11}$$

subject to two constraints. The equilibrium strategy profile in the induced game must satisfy the ensuing *interim* incentive compatibility and *ex post* individual rationality constraints.

Given the prior beliefs of players, the expected utility for type t_i of player i from participating in the induced game G and truthfully reporting $m_i = t_i$ is

$$U_i(t_i|t_i) = \int_0^1 q(t_i,t_j)x_i(t_i,t_j) + (1 - q(t_i,t_j))w_i(t_i,t_j)\, dt_j$$

and the expected utility from falsely reporting type $m_i' \neq t_i$ is

$$U_i(m_i'|t_i) = \int_0^1 q(m_i', t_j)x_i(m_i', t_j) + (1 - q(m_i', t_j))w_i(t_i, t_j)\,dt_j.$$

IC constraint states that for any $t_i, m_i' \in M$

$$U_i(t_i|t_i) \geq U_i(m_i'|t_i) \quad \textit{and} \quad U_i(m_i'|m_i') \geq U_i(t_i|m_i'). \tag{2.12}$$

It must be optimal in the equilibrium of G to accept all peaceful splits proposed. Recall that if type profile is (t_1, t_2) then the war payoff of player i is $w_i(t_i, t_j) = p(t_i, t_j)\theta(t_i, t_j)$. Given that messages are public and truthfully reveal types, the ex post IR constraints are

$$x(m_1, m_2) \geq w_1(m_1, m_2) \quad \textit{and} \quad 1 - x(m_1, m_2) \geq w_2(m_1, m_2) \tag{2.13}$$

for all m_1, m_2.

Proposition 4 *Let the direct peace talk mechanism be incentive-compatible. Then the probability of recommendation of peace inducing split is not constant in reported types.*

Proof: Let the direct mechanism with decision rule $f(m_1, m_2) = (x(m_1, m_2), q(m_1, m_2))$ be incentive-compatible. Let $m_1, m_1' \in [0,1]$ and $m_1 \neq m_1'$. The *ex ante* IC constraint (2.12) for player 1 of type m_1 yields

$$\int_0^1 q(m_1, m_2)x(m_1, m_2)\,dm_2 + \int_0^1 (1 - q(m_1, m_2))w_1(m_1, m_2)\,dm_2 \geq \tag{2.14}$$

$$\geq \int_0^1 q(m_1', m_2)x(m_1', m_2)\,dm_2 + \int_0^1 (1 - q(m_1', m_2))w_1(m_1, m_2)\,dm_2.$$

For player 1 of type m_1' the *ex ante* IC constraint yields

$$\int_0^1 q(m_1', m_2)x(m_1', m_2)\,dm_2 + \int_0^1 (1 - q(m_1', m_2))w_1(m_1', m_2)\,dm_2 \geq \tag{2.15}$$

$$\geq \int_0^1 q(m_1, m_2)x(m_1, m_2)\,dm_2 + \int_0^1 (1 - q(m_1, m_2))w_1(m_1', m_2)\,dm_2.$$

Assume that $q(m_1', m_2) = q(m_1, m_2)$ for any $m_2 \in M$. Then inequality (2.14) yields that $\exists m_2^* \in M$ such that

$$x(m_1, m_2^*) \geq x(m_1', m_2^*),$$

while inequality (2.15) yields that $\exists m_2^{**} \in M$ such that

$$x(m_1', m_2^{**}) \geq x(m_1, m_2^{**})$$

holds. As m_1, m_1' are arbitrary, it implies $x(m_1, m_2) = x(m_2)$. Applying the same reasoning to player 2 we obtain $x(m_1, m_2) = x(m_1)$. Hence, $x(m_1, m_2) = const = K$. That is, all types expect the same share in the equilibrium without war. The *ex ante* IR constraint for the highest type of player 1 states

$$\int_0^1 x(1, m_2)\, dm_2 = K \geq \int_0^1 p\theta(1, m_2)\, dm_2 = p\int_0^1 (1 - \alpha m_2)\, dm_2 = \tag{2.16}$$

$$= p(1 - \frac{\alpha}{2}) > \frac{1}{2}.$$

Similarly, the *ex ante* IR constraint for the highest type of player 2 states

$$\int_0^1 x(m_1, 1)\, dm_1 = K \geq \int_0^1 p\theta(m_1, 1)\, dm_1 = p\int_0^1 (1 - \alpha m_1)\, dm_1 = \tag{2.17}$$

$$= p(1 - \frac{\alpha}{2}) > \frac{1}{2}.$$

Taking expectations we obtain the same constant

$$\int_0^1 \int_0^1 x(m_1, m_2)\, dm_2\, dm_1 = K.$$

Therefore, the IR constraint is violated because

$$K + K \geq 2p(1 - \frac{\alpha}{2}) > 1,$$

a contadiction. □

2.3.2 Peace talk game with partial revelation

In this subsection we consider the optimal direct mechanism that allows for partial revelation of types. Truthful estimates of player's type certified by experts can be more or less informative. Precise estimates might be available if information can be acquired costless and covertly. Otherwise, it might not be an equilibrium strategy for any of players to allow assessment of his own type as accurate as possible. We assume that for each player i experts observe only an interval that contains t_i. A partial type may be any subset of $T = [0, 1]$. In general, partial types may be overlapping and they have to be exhaustive in order to define incentive compatibility of a mechanism.

Definition 4 *Let $\{T_1, T_2, ..., T_n\}$ be a finite set of intervals such that $\cup_{j=1}^{n} T_i =$ $[0,1]$. We say that a report T_i of player j is truthful if $t_j \in T_i$, that is, if the partial type T_i contains the player' s true type t_j.*

We will design a mechanism that reaches socially efficient outcome when the partial types are the least informative. We consider a mechanism in which messages reveal the type of the sender with precision up to the division of the type space to two intersecting subintervals $[0,1/2]$ and $[1/2,1]$. Without loss of generality, we restrict the message space to the set $M = [h,l]$. A truthful strategy for player i in the induced game is to report type l if $t_i \leq 1/2$ and h if $t_i \geq 1/2$. The symmetry of recommended splits $x(m_1,m_2)$ and probabilities $q(m_1,m_2)$ across players yields

$$x(l,l) = x(h,h) = 1/2 \quad and \quad x \equiv x(h,l) = 1 - x(l,h).$$

The three unknown probabilities for the mechanism are

$$q_l \equiv q(l,l), \quad q_h \equiv q(h,h), \quad q_m \equiv q(h,l) = q(l,h).$$

There are many pooling equilibria of the game $(M,(x,q))$. In order to construct one of them, assume that one of the players sends a message h with probability one. Assume that in the case of messages profile (h,h) the rule offers equal shares with probability $q(h,h) = 1$. Assume that for messages profile (h,l) the rule offers $\pi(1/2)$ to the player who sends a message l. Since each low type prefers to send a message h it means that only message h is sent by any type of each player in the Bayesian Nash equilibrium of the induced game. In each pooling equilibrium of the game induced by the mechanism the *ex ante* probability that the peace recommendation will be accepted by both players is the same as the probability of peace in the split proposal game without communication.

We will show that there exists a separating equilibrium of the game induced by $(M,(x,q))$ in which each partial type sends different message. In a Bayesian Nash equilibrium of the induced game, the peace maximizing splits and probabilities should solve the problem

$$max_{x,q_l,q_h,q_m}\left\{\frac{1}{4}q_l+\frac{1}{2}q_m+\frac{1}{4}q_h\right\}, \tag{2.18}$$

that corresponds to problem (2.11) and is subject to the probability constraints

$$0\le q_l\le 1,\ 0\le q_m\le 1,\ 0\le q_h\le 1 \tag{2.19}$$

and the following IC and IR constraints.

In a truthful mechanism messages reveal types. The *interim* expected prize in a war with reported low type is

$$\int_0^{1/2}(1-\alpha t_i m_2)\,dm_2=\frac{1}{2}-\frac{\alpha t_i}{8} \tag{2.20}$$

while the expected prize in a war with reported high type is

$$\int_{1/2}^{1}(1-\alpha t_i m_2)\,dm_2=\frac{1}{2}-\frac{3\alpha t_i}{8}. \tag{2.21}$$

Then the expected payoff of a high type from waging a war with reported low type is

$$p\int_0^{1/2}(1-\alpha t_i m_2)\,dm_2=p(\frac{1}{2}-\frac{\alpha t_i}{8}).$$

We require that the share x of a reported high type makes war against a self-reported low type unprofitable. Hence, we require

$$x\ge p(\frac{1}{2}-\frac{\alpha t_i}{8})$$

for any $t_i>1/2$. Then the *ex post* IR constraint for the high type share states

$$x\ge p(\frac{1}{2}-\frac{\alpha}{16}). \tag{2.22}$$

Similarly, the share $1-x$ of the low type should make it unprofitable to wage a war with reported high type. The expected payoff of a low type from waging a war with reported high type is

$$(1-p)\int_{1/2}^{1}(1-\alpha t_i m_2)\,dm_2)=(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8}).$$

Hence, we require

$$1-x \geq (1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})$$

for any $t_i \leq 1/2$. Then the *ex post* IR constraint for the low type share states

$$1-x \geq \frac{1}{2}(1-p). \tag{2.23}$$

Clearly, condition (2.2) implies that inequality $x = p(\frac{1}{2}-\frac{\alpha}{16}) > 1-x = \frac{1}{2}(1-p)$ is satisfied.

In the construction of the *interim* IC constraints we assume that misreporting is never followed by a failure to comply with recommendation of the public randomization device. We check later in the section that the solution of the program gives no incentive for players to deviate by waging a war after misreporting.

Denote by $\pi_l(t_i)$ the expected payoff of player i of type $t_i \leq 1/2$ from a war with low type. Denote by $\pi_h(t_i)$ the expected payoff of player i of type $t_i \geq 1/2$ from a war with high type.

Then *interim* IC constraint for a player i of type $t_i \leq 1/2$ is

$$\frac{1}{2}(\frac{q_l}{2}+(1-q_l)\pi_l(t_i))+\frac{1}{2}(q_m(1-x)+(1-q_m)(1-p)\int_{1/2}^{1}(1-\alpha t_i m_2)\,dm_2) \geq$$

$$\frac{1}{2}(q_m x+(1-q_m)\pi_l(t_i))+\frac{1}{2}(\frac{q_h}{2}+(1-q_h)(1-p)\int_{1/2}^{1}(1-\alpha t_i m_2)\,dm_2),$$

equivalent to

$$\frac{q_l}{2}-q_l\pi_l(t_i)+q_m(1-x)-q_m(1-p)\int_{1/2}^{1}(1-\alpha t_i m_2)\,dm_2 \geq \tag{2.24}$$

$$q_m x-q_m\pi_l(t_i)+\frac{q_h}{2}-q_h(1-p)\int_{1/2}^{1}(1-\alpha t_i m_2)\,dm_2$$

for all $t_i \leq 1/2$. The LHS is the expected payoff from sending a message l while the RHS is the expected payoff from exaggerating strength.

Similarly, the *interim* IC constraint for a player i of type $t_i > 1/2$ is

$$\frac{1}{2}(q_m x+(1-q_m)p\int_{0}^{1/2}(1-\alpha t_i m_2)\,dm_2)+\frac{1}{2}(\frac{q_h}{2}+(1-q_h)\pi_h(t_i)) \geq$$

$$\frac{1}{2}(\frac{q_l}{2}+(1-q_l)p\int_{0}^{1/2}(1-\alpha t_i m_2)\,dm_2)+\frac{1}{2}(q_m(1-x)+(1-q_m)\pi_h(t_i)),$$

equivalent to

$$q_m x - q_m p\int_0^{1/2}(1-\alpha t_i m_2)\,dm_2 + \frac{q_h}{2} - q_h\pi_h(t_i) \geq \tag{2.25}$$

$$\frac{q_l}{2} - q_l p\int_0^{1/2}(1-\alpha t_i m_2)\,dm_2 + q_m(1-x) - q_m\pi_h(t_i)$$

for all $t_i > 1/2$. The LHS is the expected payoff from sending a message h while the RHS is the expected payoff from hiding strength.

By substituting prize values (2.20) and (2.21) in inequalities (2.25) and (2.24) respectively, we obtain IC constraints

$$\frac{q_l}{2} - q_l\pi_l(t_i) + q_m(1-x) - q_m(1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8}) \geq \tag{2.26}$$

$$\geq q_m x - q_m\pi_l(t_i) + \frac{q_h}{2} - q_h(1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8})$$

for all $t_i \leq 1/2$ and

$$q_m x - q_m p(\frac{1}{2} - \frac{\alpha t_i}{8}) + \frac{q_h}{2} - q_h\pi_h(t_i) \geq \frac{q_l}{2} - q_l p(\frac{1}{2} - \frac{\alpha t_i}{8}) + q_m(1-x) - q_m\pi_h(t_i) \tag{2.27}$$

for all $t_i > 1/2$.

We will be using the following lemma.

Lemma 2 *For parameters satisfying condition* (2.2)*, expected payoff* $\pi_l(t_i)$ *of a low type* $t_i \leq 1/2$ *from a war with low type satisfies*

$$\pi_l(0) = \frac{1}{2}(1-p) \leq \pi_l(t_i) \leq \pi_l(\frac{1}{2}) = \frac{1}{2}p(1-\frac{\alpha}{8}) < 1/2. \tag{2.28}$$

and expected payoff $\pi_h(t_i)$ *of a high type* $t_i \geq 1/2$ *from a war with high type satisfies*

$$\pi_h(1/2) = \frac{1}{2}(1-p)(1-\frac{3}{8}\alpha) \leq \pi_h(t_i) \leq \pi_h(1) = \frac{1}{2}p(1-\frac{3}{4}\alpha) < 1/2. \tag{2.29}$$

Proof: Expected payoff of type $t_i \leq 1/2$ from a war with type $t_j \leq 1/2$ is

$$\pi_l(t_i) = \int_0^{1/2} p(t_i,t_j)\theta(t_i,t_j)\,dt_j = \int_0^{t_i} p(1-\alpha t_1 t_2)\,dt_2 + \int_{t_i}^{1/2}(1-p)(1-\alpha t_i t_j)\,dt_j =$$

$$= \frac{1-p}{2} + \left(2p - 1 - \frac{\alpha(1-p)}{8}\right)t_i + \left(\frac{\alpha(1-p)}{2} - \frac{\alpha p}{2}\right)t_i^3. \tag{2.30}$$

It is easy to check that for $t_i \in [0,1/2]$ and parameter values satisfying condition (2.2)

$$\frac{\partial \pi_l(t_i)}{\partial t_i} > 0.$$

Hence, $\pi_l(t_i)$ is increasing in t_i for $t_i \in [0,1/2]$. Expected payoff of a higher type $t_i > 1/2$ from a war with higher type $t_j > 1/2$ is

$$\pi_h(t_i) = \int_{1/2}^{1} p(t_i,t_j)\theta(t_i,t_j)\,dt_j = \int_{1/2}^{t_i} p(1-\alpha t_i t_j)\,dt_j + \int_{t_i}^{1} (1-p)(1-\alpha t_i t_j)\,dt_j =$$

$$= 1 - \frac{3p}{2} + \left(2p - 1 - \frac{\alpha}{2} + \frac{5\alpha p}{8}\right) t_i + \frac{\alpha(1-2p)}{2} t_i^3. \tag{2.31}$$

It is easy to check that for $t_i \in [1/2,1]$ and parameters values satisfying condition (2.2)

$$\frac{\partial \pi_h(t_i)}{\partial t_i} > 0.$$

Hence, $\pi_h(t_i)$ is increasing in t_i for $t_i \in [1/2,1]$. $\square$

Consider the game G induced by the mechanism $g = ([h,l] \times [h,l], f)$ where the decision rule f is determined by split function $x(m_1,m_2)$ and war probability function $q(m_1,m_2)$ that satisfy IC constraints (2.26) and (2.27), probability constraints (2.19), and solve problem (2.18).

Proposition 5 *There is an unique best separating equilibrium of the game G. The ex ante probability of peace in this equilibrium equals*

$$\mathscr{P} = \frac{1}{2} + \frac{8 - 16p + 4\alpha p - 3\alpha}{2(2\alpha p - 3a - 8)}.$$

Proof: The proof is by construction. We calculate parameters of the direct mechanism and we show that both types have no incentive to deviate from recommendations of the public randomization device.

We rearrange the IC constraint (2.26) for the low type and we consider a relaxed problem: maximizing (2.18) subject to the high type *ex post* IR constraint

$$x \geq p\left(\frac{1}{2} - \frac{\alpha}{16}\right), \tag{2.32}$$

the probability constraints

$$q_l \leq 1,\ 0 \leq q_m \leq 1,\ q_h \leq 1, \tag{2.33}$$

and the low type *ex ante* IC constraint

$$q_l\left(\frac{1}{2}-\pi_l(t_i)\right)\geq q_m\left(2x-1-\pi_l(t_i)+(1-p)\left(\frac{1}{2}-\frac{3\alpha t_i}{8}\right)\right)+\qquad(2.34)$$
$$+q_h\left(\frac{1}{2}-(1-p)\left(\frac{1}{2}-\frac{3\alpha t_i}{8}\right)\right).$$

1. By Lemma (2) inequality $\frac{1}{2}-\pi_l(t_i)>0$ holds for any $t_i\leq 1/2$. Then setting $q_l=1$ maximizes the LHS of (2.34) and does not affect the RHS. Simultaneously, it does not affect the high type *ex post* IR constraint (2.32).

2. As $\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})>0$ for any $t_i\leq 1/2$, it follows that at the maximal feasible value of q_h the IC constraint (2.34) binds for some $t_i^*\leq 1/2$. In the light of step (1) we rewrite the IC constraint for the low type as

$$\frac{1}{2}\geq(1-q_m)\pi_l(t_i)+(q_h-q_m)(1-p)\frac{3\alpha t_i}{8}-(q_h-q_m)\frac{(1-p)}{2}+\qquad(2.35)$$
$$+q_m(2x-1)+q_h\frac{1}{2}.$$

Hence, t_i^* maximizes the value of $(1-q_m)\pi_l(t_i)+(q_h-q_m)(1-p)\frac{3\alpha t_i}{8}$.

3. We want to show that the high type *ex post* IR constraint (2.32) binds. Suppose that it is slack, that is, $x>p(\frac{1}{2}-\frac{\alpha}{16})$. Then it is possible to reduce x without violating the IC constraint (2.35) because x appears in the RHS of (2.35) with coefficient $2q_m\geq 0$. It makes the constraint (2.35) slack also for t_i^*, a contradiction with (2). Therefore, the high type *ex post* IR constraint (2.32) binds.

4. Steps (1) and (3) yield

$$x=p\left(\frac{1}{2}-\frac{\alpha}{16}\right)\quad and\quad q_l=1.\qquad(2.36)$$

We want to show that $q_h\geq q_m$. In the light of equalities (2.36) the constraint (2.35) which is binding for $t_i=t_i^*$ becomes

$$\frac{1}{2}-\pi_l(t_i^*)=q_m\left(p(1-\frac{\alpha}{8})-1-\pi_l(t_i^*)+(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})\right)+$$
$$+q_h\left(\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})\right).\qquad(2.37)$$

As $\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})>0$,

$$q_h = \frac{\frac{1}{2}-\pi_l(t_i^*)}{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})} + q_m \frac{1-p(1-\frac{\alpha}{8})+\pi_l(t_i^*)-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})}{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})} \tag{2.38}$$

is well defined. We rearrange (2.38) as

$$\begin{aligned} q_h = q_m \frac{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})}{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})} + \frac{q_m(1-p(1-\frac{\alpha}{8})-\frac{1}{2}+\pi_l(t_i^*))+\frac{1}{2}-\pi_l(t_i^*)}{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})} = \\ = q_m + \frac{q_m(1-p(1-\frac{\alpha}{8}))+(1-q_m)(\frac{1}{2}-\pi_l(t_i^*))}{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})}. \end{aligned} \tag{2.39}$$

By lemma (2) inequality $\frac{1}{2}-\pi_l(t_i^*)>0$ holds. Hence, the ratio on the RHS of (2.39) is positive for any $q_m \le 1$. Therefore, $q_h \ge q_m$.

5. We want to show that the solution of the relaxed problem is $q_l = q_h = 1$ and $q_m = \frac{8-16p+4\alpha p-3\alpha}{2\alpha p-3a-8}$.

   In the light of step (1) we rewrite the IC constraint (2.34) as

$$\frac{1}{2} \ge (1-q_m)\pi_l(t_i)+(q_h-q_m)(1-p)\frac{3\alpha t_i}{8}-(q_h-q_m)\frac{(1-p)}{2}+q_m(2x-1)+q_h\frac{1}{2}. \tag{2.40}$$

   As $q_h-q_m \ge 0$ and $1-q_m \ge 0$, lemma (2) implies that the value of $(1-q_m)\pi_l(t_i)+(q_h-q_m)(1-p)\frac{3\alpha t_i}{8}$ is maximal for $t_i=1/2$. Therefore, the IC constraint (2.40) binds for $t_i^*=1/2$. Then constraint

$$\begin{aligned} \frac{1}{2} \ge (1-q_m)\frac{1}{2}p(1-\frac{\alpha}{8})+(q_h-q_m)(1-p)\frac{3\alpha}{16}-(q_h-q_m)\frac{(1-p)}{2}+ \\ +q_m(2x-1)+q_h\frac{1}{2} \end{aligned} \tag{2.41}$$

   binds.

   We substitute x in the binding constraint (2.41) and we obtain

$$\begin{aligned} 1 = (1-q_m)p(1-\frac{\alpha}{8})+(q_h-q_m)(1-p)\frac{3\alpha}{8}-(q_h-q_m)(1-p)+ \\ +q_m(p(2-\frac{\alpha}{4})-2)+q_h. \end{aligned} \tag{2.42}$$

Clearly, $2\alpha p - 3a - 8 < 0$ for any feasible values of α and p. Hence,

$$q_m = q_h \frac{3\alpha p - 8p - 3\alpha}{2\alpha p - 3a - 8} + \frac{8 - 8p + \alpha p}{2\alpha p - 3a - 8} \tag{2.43}$$

is well defined. In the light of step (1) we simplify the objective function (2.18) and maximize

$$max_{q_h, q_m} \{2q_m + q_h\}. \tag{2.44}$$

Substituting q_m by the RHS of (2.43) we maximize expression

$$W = \frac{2(8 - 8p + \alpha p)}{2\alpha p - 3a - 8} + q_h \left(1 + \frac{2(3\alpha p - 8p - 3\alpha)}{2\alpha p - 3a - 8}\right).$$

We note that coefficient of q_h is positive and the maximization of W requires maximization of q_h. However, the value of q_h is constrained by inequality in (2.33). Setting $q_h = 1$ and solving for q_m in (2.43) yields

$$q_m = \frac{3\alpha p - 8p - 3\alpha}{2\alpha p - 3a - 8} + \frac{8 - 8p + \alpha p}{2\alpha p - 3a - 8} = \frac{8 - 16p + 4\alpha p - 3\alpha}{2\alpha p - 3a - 8}. \tag{2.45}$$

It is easy to check that $0 < q_m \leq 1$ for any feasible values of p and α. Indeed, $2\alpha p - 3a - 8 < 0$ and inequality $8 - 16p + 4\alpha p - 3\alpha < 0$ holds for any values of parameters α and p that satisfy condition (2.2). The RHS of (2.45) achieves its maximal value of $\frac{8-\alpha}{8+\alpha} \leq 1$ for $p = 1$. Therefore, the solution $q_l = q_h = 1$ and $q_m = \frac{8-16p+4\alpha p-3\alpha}{2\alpha p-3a-8}$ is admissible.

6. We want to show that solution constructed in step (5) satisfies all constraints of the initial problem. The *ex post* IR constraint(2.23) for the low type share is trivially satisfied when $x = p(\frac{1}{2} - \frac{\alpha}{16})$, as $1 - x = 1 - p(\frac{1}{2} - \frac{\alpha}{16}) \geq \frac{1}{2}(1 - p)$. By substituting $q_l = q_h = 1$ the high-type *ex ante* IC constraint (2.27) becomes

$$\frac{1}{2} - \pi_h(t_i) + q_m x - q_m p(\frac{1}{2} - \frac{\alpha t_i}{8}) \geq q_m(1 - x) - q_m \pi_h(t_i) + \frac{1}{2} - p(\frac{1}{2} - \frac{\alpha t_i}{8}) \tag{2.46}$$

We will show that inequality (2.46) is satisfied for any $t_i \geq 1/2$.

Recall that by inequality (2.29) of lemma (2) inequality

$$p(\frac{1}{2} - \frac{\alpha t_i}{8}) \geq \pi_h(t_i) \tag{2.47}$$

holds for $t_i \geq 1/2$. Then

$$(1-q_m)p(\frac{1}{2}-\frac{\alpha t_i}{8}) \geq (1-q_m)\pi_h(t_i) \tag{2.48}$$

because $q_m < 1$. We rewrite (2.48) as

$$-\pi_h(t_i) - q_m p(\frac{1}{2}-\frac{\alpha t_i}{8}) \geq -q_m\pi_h(t_i) - p(\frac{1}{2}-\frac{\alpha t_i}{8}). \tag{2.49}$$

We note that summing inequality (2.49) with inequality $q_m x > q_m(1-x)$, that holds because $x > 1-x$, we obtain (2.46), which had to be proved. The probability constraints (2.19) are obviously satisfied.

By lemma (2) the value of split $x = p(\frac{1}{2}-\frac{\alpha}{16})$ implies that neither the low nor the high type has any incentive to deviate by waging a war after misreporting, learning the type of the opponent and receiving a peaceful recommendation by the mechanism. Hence, truthfully reporting type and following the recommendation is a Bayesian-Nash equilibrium in the game induced by the mechanism. This equilibrium is unique up to the strategy of the player with relative strength $t_i = 1/2$ who can randomize between reporting h and l. Therefore, in the best separating equilibrium of the peace talk game the *ex ante* probability of peace is

$$\mathscr{P} = \frac{1}{4} + \frac{8-16p+4\alpha p-3\alpha}{2(2\alpha p-3a-8)} + \frac{1}{4}.$$

□

We note that *ex ante* probability of peace in the separating equilibrium of the peace talk game satisfies $\mathscr{P} > \frac{1}{2}$ for any parameters values. Therefore, by inequality (2.8), for parameter values satisfying condition (2.2) a peace talk reduces the *ex ante* probability of conflict compared to all equilibria of the split offer game with no communication.

In the ex post stage, as a result of updated beliefs about the opponent's war capacity, parties might not follow a war recommendation and seek bilaterally another peaceful mechanism. However, the lack of commitment to fight after a war recommendation can only increase the ex ante probability of peace in the case of communication.

2.4 Probability of peace in mediation games

We consider a mediated communication game studied by Horner, Morelli and Squintani (2010) in [11]. Typically, the mediator is a third party that has no private information but helps others to reach an agreement. The mediator who evaluates the case is expected to adopt a neutral stance in his advisory role. Nevertheless, a typical mediator has his own agenda. The fundamental assumption in our model is that the objective of a mediator is to maximise the *ex ante* probability of peaceful resolution of the conflict and this objective is common knowledge. The mediator cannot enforce his recommendations. We assume that the mediator can fully commit to the outcome induced by the mechanism and that channels of communication of the privately informed players with the mediator are perfect and immune to disclosure of confidential information.

The timing of the mediation game is as follows. After observing his own type, each player sends simultaneously a message. The mediator receives messages and offers a mechanism. Parties know how their messages are to be used. The mediator commits himself to a policy rule that maximises the probability of peace. By the Revelation Principle for Bayesian equilibrium we restrict our consideration to direct mechanisms offered by the mediator.

In contrast to the peace talk game, messages sent by parties to the mediator are non publicly observed. The mediator randomly selects a split from a full menu, but this recommendation is made after learning the messages profile m. Following the model by Horner, Morelli and Squintani (2010) presented in [11], we assume that a split $(x, 1-x)$ is recommended according to some cumulative distribution function $F(x|m)$ over a set of possibilities where there is only one recommendation in the support of F leading to war, the split $(0,1)$. After receiving a recommendation opponents play an agreement game with the proposed split.

S. Baliga and T. Sjostrom (2011) express in [1] the view that 'the mechanisms are clearly not meant to be descriptive of real-world institutions. For example, they typically require the agents to report "all they know" before any decision is reached, an extreme form of centralized decision making hardly ever encountered in the real world.' Several papers study the impact of the alignment of preferences of the sender and the receiver on the amount of information communicated in equilibrium (see [9] and [10] by Kamenica and Gentzkow (2009,2011)). The authors show that more aligned preferences can make the optimal signal either more or less informative depending on the default action of the receiver. In general, parties have a control over the precision of the information shared with the mediator. We relax the assumption of full disclosure of private information to the mediator. We consider mediation games in which players can only partially reveal their types.

2.4.1 Revelation of information through the choice of recommendation

We consider equilibria in which messages sent to the mediator reveal the type of the sender with precision up to the division of the type space to subintervals $[0,1/2]$ and $[1/2,1]$. Given the messages profile $m=(m_1,m_2)$ where $m_i \in \{h,l\}$, the mediator recommends a split chosen from a set of possibilities. Following the model of Horner et al., we consider only direct mechanisms with discrete and symmetric randomisation function F.

The mediator has to decide how informative should be his recommendation for players. Players can choose to go to war at any time. They use information learned from the mediator's recommendation for updating their belief about the type of opponent. Hence, in the construction of the menu of lotteries $F(x|m)$, the mediator should avoid releasing to players unnecessary information about opponent's type. Following the recommendation should be optimal given players' types and the updated beliefs about the opponent's type, where beliefs are consistent with Bayes' rule.

In general, the mediator can assign a positive probability to a number of peaceful splits. Following the analysis in [11], we restrict our attention to mediators whose set of recommendations $R=\{(1/2,1/2),(x,1-x),(1-x,x),(0,1)\}$, where $x>1-x$, contains only four elements. Clearly, the split $(0,1)$ induces war with probability 1. We consider a distribution function $F:\{h,l\}\times\{h,l\}\to[0,1]^4$ characterized by five probabilities. If recommendations $(x,1-x)$ and $(x,1-x)$ are made only for messages profiles (l,h) and (h,l) then with probability $1/2$ in the truthful Bayesian equilibrium players learn *ex post* the partial type of the opponent. Such a mechanism can not be incentive compatible for the partial type l. Hence, an incentive compatible mediation programme should offer with positive probability recommendations $(x,1-x)$ and $(x,1-x)$ to players with the same partial type. In the following paragraphs we study properties of mediation programmes that recommend with positive probability uneven split of the prize not only in the case of heterogeneous messages profiles but also in the case of (l,l) profile or in the case of (h,h) profile.

2.4.2 Mediation programme mixing for low types

In this section we consider the mediation programme given by

$$F(h,h) = (q_h, 0, 0, 1-q_h), \quad F(h,l) = (q_m, p_m, 0, 1-q_m-p_m), \tag{2.50}$$

$$F(l,l) = (q_l, p_l, p_l, 1-q_l-2p_l).$$

By symmetry $F(l,h) = (q_m, 0, p_m, 1-q_m-p_m)$. For this type of programmes the mediator chooses the value of x and probabilities that maximize the probability of peace

$$max_{x,q_l,q_h,q_m,p_l,p_m} \left\{ \frac{1}{4}q_h + \frac{1}{2}(q_m+p_m) + \frac{1}{4}(q_l+2p_l) \right\}, \tag{2.51}$$

subject to the probability constraints

$$0 \le q_l \le 1,\ 0 \le q_m \le 1,\ 0 \le q_h \le 1, 0 \le p_l \le 1,\ 0 \le p_m \le 1. \tag{2.52}$$

The objective function is also subject to IC and IR constraints that follow.

First, we consider two *ex post* IR constraints for the high type share. A reported high type is recommended x with probability p_m. After this recommendation he updates his belief about the type of the opponent and $Pr[m_2 = l|\ x \wedge m_1 = h] = 1$. We require that war is unprofitable *ex post* for a reported high type who receives a peaceful recommendation x. Hence, an *ex post* IR constraint for the high type share states

$$xp_m \ge p\left(\frac{1}{2} - \frac{\alpha t_i}{8}\right)$$

for any $t_i \ge 1/2$, which is equivalent to

$$xp_m \ge p\left(\frac{1}{2} - \frac{\alpha}{16}\right). \tag{2.53}$$

A reported high type is proposed $1/2$ with probability $q_h + q_m$. After this recommendation he updates his belief about the type of the opponent as $Pr[1/2\ |m_1 = h] = P[m_2 = h|\ 1/2 \wedge m_1 = h] + P[m_2 = l\ |\ 1/2 \wedge m_1 = h]$. Hence, another *ex post* IR constraint for the high type share states

$$\frac{1}{2}(q_h+q_m) \ge q_h\pi_h(t_i) + q_m p\left(\frac{1}{2} - \frac{\alpha t_i}{8}\right)$$

for any $t_i \geq 1/2$. Clearly, this inequality holds for any q_h and q_m which satisfy the probability constraints because $\pi_h(t_i) < \frac{1}{2}$ by lemma (2).

Next, we consider three *ex post* IR constraints for the low type share. A reported low type is recommended x with probability p_l. After this recommendation he updates his belief about the type of the opponent and $Pr[m_2 = l|\ x \wedge m_1 = l] = 1$. Hence, war is unprofitable *ex post* for a low type who receives a recommendation x if

$$xp_l \geq \pi_l(t_i)$$

for any $t_i \leq 1/2$. Therefore, by lemma (2) an IR constraint for a low type share states

$$xp_l \geq p\left(\frac{1}{2} - \frac{\alpha}{16}\right). \tag{2.54}$$

A reported low type is recommended $1-x$ with probability $Pr[1-x|m_1 = l] = P[m_2 = h|\ 1-x \wedge m_1 = l] + P[m_2 = l|\ 1-x \wedge m_1 = l] = p_m + p_l$. We require that war is unprofitable *ex post* for a low type who receives a peaceful recommendation $1-x$. Hence, another *ex post* IR constraint for the low type share states

$$(p_m + p_l)(1-x) \geq p_m(1-p)\left(\frac{1}{2} - \frac{3\alpha t_i}{8}\right) + p_l\pi_l(t_i) \tag{2.55}$$

for any $t_i \leq 1/2$.

A reported low type is proposed $1/2$ with probability $Pr[1/2|m_1 = l] = P[m_2 = h|\ 1/2 \wedge m_1 = l] + P[m_2 = l|\ 1/2 \wedge m_1 = l] = q_m + q_l$. We require that war is unprofitable *ex post* for a low type who receives a peaceful recommendation $1/2$. Hence, another *ex post* IR constraint for the low type share states

$$\frac{1}{2}(q_m + q_l) \geq q_m(1-p)\left(\frac{1}{2} - \frac{3\alpha t_i}{8}\right) + q_l\pi_l(t_i)$$

for any $t_i \leq 1/2$. This IR constraint is clearly satisfied because $\pi_l(t_i) < \frac{1}{2}$ by lemma (2).

In general, the mediator is not endowed with any power to enforce compliance with his recommendations. However, in the construction of the *interim* IC constraints for the mediation game we assume that misreporting is never followed by a failure to comply with recommendation, similarly to the IC constraints for the peace talk game. We will check later in the section if the solution of the program, in which the recommendation is not enforced, provides some incentive for players to deviate by waging a war after misreporting.

Truthful reporting by a high type requires that the following *interim* IC constraint

$$\frac{1}{2}((1-q_h)\pi_h(t_i)+q_h\frac{1}{2})+\frac{1}{2}((1-q_m-p_m)p(\frac{1}{2}-\frac{\alpha t_i}{8})+q_m\frac{1}{2}+p_m x)\geq \tag{2.56}$$

$$\geq\frac{1}{2}\left((1-q_m-p_m)\pi_h(t_i)+q_m\frac{1}{2}+p_m(1-x)\right)+$$

$$+\frac{1}{2}\left((1-q_l-2p_l)p(\frac{1}{2}-\frac{\alpha t_i}{8})+q_l\frac{1}{2}+p_l x+p_l(1-x)\right)$$

is satisfied for all $t_i\geq 1/2$.

Truthful reporting by a low type requires that the following *interim* IC constraint

$$\frac{1}{2}\left((1-q_m-p_m)(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})+q_m\frac{1}{2}+p_m(1-x)\right)+ \tag{2.57}$$

$$+\frac{1}{2}\left((1-q_l-2p_l)\pi_l(t_i)+q_l\frac{1}{2}+p_l x+p_l(1-x)\right)\geq$$

$$\geq\frac{1}{2}((1-q_h)(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})+q_h\frac{1}{2})+\frac{1}{2}((1-q_m-p_m)\pi_l(t_i)+q_m\frac{1}{2}+p_m x)$$

is satisfied for all $t_i\leq 1/2$.

Consider the game G induced by the mediation programme (2.50) where the decision rule satisfies IC constraints (2.57) and (2.56), probability constraints (2.52), and solves the problem (2.51).

Proposition 6 *There is an unique best separating equilibrium of the game G induced by the mediation programme* (2.50). *The ex ante probability of peace in this equilibrium equals*

$$\mathscr{P}_{med}=\frac{1}{2}+\frac{32p-24+3\alpha-16\alpha p}{2(8+3\alpha)}.$$

Proof: The proof is by construction. We calculate parameters of the optimal mediation programme and we show that both types have no incentive to deviate from recommendations of the mediator.

For the sake of convenience we change the parametrisation of the relaxed program by introducing four variables. We set $s_h\equiv q_h$, $s_m\equiv q_m+p_m$ and $s_l\equiv$

$q_l + 2p_l$. Additionally, we denote by

$$b = \frac{1}{2}q_m + xp_m$$

the expected payoff of reported high type in the case of peaceful agreement with a low type. We note that the expected payoff of reported low type in the case of peaceful agreement with a high type is

$$\frac{1}{2}q_m + (1-x)p_m = s_m - b$$

and the expected payoff of reported low type in the case of peaceful agreement with a low type is

$$\frac{1}{2}q_l + xp_l + (1-x)p_l = \frac{1}{2}q_l + p_l = \frac{1}{2}s_l.$$

We solve the following relaxed problem:

$$max_{b,s_l,s_m,s_h} \left\{ \frac{1}{4}s_h + \frac{1}{2}s_m + \frac{1}{4}s_l \right\}, \tag{2.58}$$

subject to the probability constraints

$$s_l \leq 1,\ 0 \leq s_m \leq 1,\ 0 \leq s_h \leq 1. \tag{2.59}$$

The *interim* IR constraint requires that, after learning its own type but before receiving any recommendation, participation in the mechanism and following all recommendations is at least as profitable for a player as going to war. The *interim* IR constraint for the high type states

$$\frac{1}{2}(s_h\frac{1}{2} + (1-s_h)\pi_h(t_i)) + \frac{1}{2}(b + (1-s_m)p(\frac{1}{2} - \frac{\alpha t_i}{8})) \geq \frac{1}{2}\pi_h(t_i) + \frac{1}{2}p(\frac{1}{2} - \frac{\alpha t_i}{8})$$

for any $t_i \geq 1/2$, which is equivalent to

$$s_h(\frac{1}{2} - \pi_h(t_i)) + b - s_m p(\frac{1}{2} - \frac{\alpha t_i}{8}) \geq 0 \tag{2.60}$$

for any $t_i \geq 1/2$.

The *interim* IC constraint for the low type states

$$\frac{1}{2}(\frac{1}{2}s_l + (1-s_l)\pi_l(t_i)) + \frac{1}{2}((s_m - b) + (1-s_m)(1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8})) \geq$$

$$\frac{1}{2}(b+(1-s_m)\pi_l(t_i))+\frac{1}{2}(s_h\frac{1}{2}+(1-s_h)(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8}))$$

for any $t_i \leq 1/2$, which is equivalent to

$$s_l(\frac{1}{2}-\pi_l(t_i))+s_m-s_m(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})\geq 2b-s_m\pi_l(t_i)+s_h(\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})) \tag{2.61}$$

for any $t_i \leq 1/2$.

1. We note that s_l appears only in right-hand side of the low type *interim* IC constraint and this right-hand side is increasing in s_l because by lemma (2) $\pi_l(t_i) \leq \frac{1}{2}$. Hence, $s_l = 1$ in the solution.

2. We note that by lemma (2) the coefficient of s_h in the right -hand side of the high type *interim* IR constraint is positive as well, and the coefficient $\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})$ in the left-hand side of the low type *interim* IC constraint is also positive. Hence, the low type *interim* IC constraint must be binding for some $t_i \leq 1/2$ in the solution of relaxed problem, otherwise we could increase s_h thus increasing the value of the objective function, without violating the high type *interim* IR constraint.

3. We note that the high type *interim* IR constraint must be binding for some $t_i \geq 1/2$ in the solution of relaxed problem, otherwise we could decrease b and make the low type *interim* IC constraint slack.

4. We would like to show that $s_h > s_m$ in the solution of relaxed problem, which implies that the high type *interim* IC constraint binds for $t_i = 1/2$. Then we show that the *interim* IR constraint for the high type binds for $t_i = 1$.

   In the light of step one, the constraint (2.61) which binds for $t_i = t_i^*$ becomes

$$\frac{1}{2}-\pi_l(t_i^*)+s_m\pi_l(t_i^*)-2b+s_m\frac{1}{2}+s_m\left(\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})\right)=$$

$$=s_h\left(\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})\right).$$

   As $\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})>0$ for any t_i, parameter

$$s_h = s_m + \frac{\frac{1}{2}-\pi_l(t_i^*)+s_m\pi_l(t_i^*)+s_m\frac{1}{2}-2b}{\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})} \tag{2.62}$$

is well defined. We would like to show that the second term in the right-hand side of (4.8) is positive. Clearly, $\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i^*}{8})>0$. It remains to show that inequality

$$\frac{1}{2}-\pi_l(t_i^*)+s_m\pi_l(t_i^*)+s_m\frac{1}{2}-2b>0 \tag{2.63}$$

holds.

The constraint (2.60) binds for some $t_i=t_i^{**}$, therefore

$$b=s_m p\left(\frac{1}{2}-\frac{\alpha t_i^{**}}{8}\right)-s_h\left(\frac{1}{2}-\pi_h(t_i^{**})\right).$$

We note that $\frac{1}{2}-\pi_h(t_i^{**})>0$ by lemma (2). Hence, $b\leq s_m p\frac{1}{2}$ for any feasible $s_h\geq 0$.

Therefore,

$$\frac{1}{2}-\pi_l(t_i^*)+s_m\pi_l(t_i^*)+s_m\frac{1}{2}-2b>\frac{1}{2}+(s_m-1)\pi_l(t_i^*)+s_m\frac{1}{2}-s_m p. \tag{2.64}$$

We note that in the solution of relaxed problem $s_m-1\leq 0$. Hence, the right-hand side of (4.10) is minimal when the value of $\pi_l(t_i^*)$ is maximal, that is, for $t_i=1/2$. Hence,

$$\frac{1}{2}+(s_m-1)\pi_l(t_i^*)+s_m\frac{1}{2}-s_m p\geq\frac{1}{2}+(s_m-1)\frac{p}{2}(1-\frac{\alpha}{8})+s_m(\frac{1}{2}-p)=$$

$$=\frac{1}{2}-\frac{p}{2}+\frac{\alpha p}{16}+s_m(\frac{p}{2}(1-\frac{\alpha}{8})+\frac{1}{2}-p)=\frac{1}{2}-\frac{p}{2}+\frac{\alpha p}{16}+s_m(\frac{1}{2}-\frac{p}{2}-\frac{\alpha p}{16})=$$

$$=g(s_m).$$

For parameter values for which $\frac{1}{2}-\frac{p}{2}-\frac{\alpha p}{16}>0$ the value of $g(s_m)$ is minimal for $s_m=0$. Then $g(0)=\frac{1}{2}-\frac{p}{2}+\frac{\alpha p}{16}>0$ for any parameters values. For parameter values for which $\frac{1}{2}-\frac{p}{2}-\frac{\alpha p}{16}<0$ the value of $g(s_m)$ is minimal for $s_m=1$. Then the minimal value of $g(s_m)$ is $g(1)=1-p>0$. Hence, for any parameters values $g(s_m)>0$ and inequality (4.9) is satisfied.

Therefore, in the light of equation (4.8) we conclude that $s_h>s_m$ in the solution of relaxed problem. In the light of step 1 we rewrite the *interim*

IC constraint for the low type (2.61) as

$$\frac{1}{2} + s_m - s_m(1-p)\frac{1}{2} \geq$$

$$\geq 2b + (1-s_m)\pi_l(t_i) + s_h\frac{1}{2} - s_h(1-p)\frac{1}{2} + (s_h - s_m)(1-p)\frac{3\alpha t_i}{8}.$$

Hence, the right hand side of this IC constraint is maximal for $t_i = 1/2$ because $1 - s_m \geq 0$, $s_h - s_m > 0$, and by lemma (2) $\pi_l(t_i)$ is maximal for $t_i = 1/2$. Therefore, the *interim* IC constraint for the low type binds for $t_i = 1/2$.

It is easy to check that for $s_h > s_m$ and $t_i \in [1/2, 1]$ the value of expression $s_m p \frac{\alpha t_i}{8} - s_h \pi_h(t_i)$ is minimal for $t_i = 1$. Hence, the IR constraint for the high type binds for $t_i = 1$. Therefore, the IR constraint

$$s_h\left(\frac{1}{2} - \pi_h(1)\right) + b - s_m p\left(\frac{1}{2} - \frac{\alpha}{8}\right) \geq 0$$

and the IC constraint

$$\frac{1}{2} - \pi_l(1/2) + s_m - s_m(1-p)\left(\frac{1}{2} - \frac{3\alpha}{16}\right) \geq$$

$$\geq 2b - s_m\pi_l(1/2) + s_h\left(\frac{1}{2} - (1-p)\left(\frac{1}{2} - \frac{3\alpha}{16}\right)\right)$$

bind.

5. We rewrite the constraints of the relaxed problem by substituting

$$b = s_m p(\frac{1}{2} - \frac{\alpha}{8}) - s_h\left(\frac{1}{2} - p(\frac{1}{2} - \frac{3\alpha}{8})\right)$$

in the IC constraint for $t_i = 1/2$

$$\frac{1}{2} - \frac{1}{2}p(1 - \frac{\alpha}{8}) + s_m - s_m(1-p)(\frac{1}{2} - \frac{3\alpha}{16}) =$$

$$= 2b - s_m\frac{1}{2}p(1 - \frac{\alpha}{8}) + s_h(\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha}{16}))$$

and we obtain

$$s_m = -\frac{8 - 8p + \alpha p}{8 + 3\alpha} - s_h\frac{16 - 3\alpha - 24p + 15\alpha p}{8 + 3\alpha} \tag{2.65}$$

In the light of step 1 we simplify the objective function (2.58) and maximize

$$max_{s_h,s_m}\left\{2s_m+s_h\right\}. \tag{2.66}$$

Substituting s_m by the RHS of (2.65) we maximize expression

$$W = const + s_h\frac{48p-24+9\alpha-30\alpha p}{8+3\alpha}.$$

We note that coefficient of s_h is positive for any values of parameters α and p that satisfy condition (2.2). Hence, the maximization of W requires maximization of s_h. However, the value of s_h is bounded by probability constraints (2.59). We distinguish two cases for values of parameters α and p.

Let $\frac{1}{6}(19-\sqrt{265}) < \alpha \leq 1/2$ or

$$0<\alpha\leq\frac{1}{6}(19-\sqrt{265})\ \wedge\ \frac{24-3\alpha}{32-16\alpha}<p\leq 1, \tag{2.67}$$

where $\frac{1}{6}(19-\sqrt{265}) \approx 0.45353$. For this range of parameters we set $s_h = 1$. We calculate s_m from (2.65) and we obtain

$$0<s_m=\frac{32p-24+3\alpha-16\alpha p}{8+3\alpha}<1. \tag{2.68}$$

Let $0<\alpha\leq\frac{1}{6}(19-\sqrt{265})\ \wedge\ 1/2<p<\frac{24-3\alpha}{32-16\alpha}$. For this range of parameters the maximal value of s_h satisfying all constraints of the relaxed problem is the value corresponding to the minimal value of s_m satisfying the high type IR constraint (2.53). Clearly, the IR constraint requires $s_m > 0$ and for this range of parameters the corresponding value of s_h and, hence, the value of W is lower.

We show in subsequence that for parameter values satisfying condition (2.67) the *ex ante* probability of peace in the Bayesian equilibrium of the induced game is lower than the *ex ante* probability of peace in the non-mediated game.

Let the solution of the initial problem (2.51) be $q_h = 1$,

$$q_m+p_m=\frac{32p-24+3\alpha-16\alpha p}{8+3\alpha}$$

and $q_l+2p_l=1$. In this case diads (h,h) and (l,l) do not fight in the Bayesian equilibrium of the induced game. We can choose the value of x in such a way that this solution does not violate the high-type IC constraint and the low type IR constraint. However, the *ex ante* probability of peace in the best separating equilibrium of the mediation game is

$$\mathscr{P}_{med} = \frac{1}{2} + \frac{32p - 24 + 3\alpha - 16\alpha p}{2(8+3\alpha)}$$

where inequality

$$\mathscr{P}_{med} < \mathscr{P}_{talk} = \frac{1}{2} + \frac{8 - 16p + 4\alpha p - 3\alpha}{2(8 - 3\alpha - 2\alpha p)}$$

holds for any feasible values of parameters p and α. □

We notice that for any feasible parameter values the *ex ante* probability of peace in the mediation game is higher than in the case of lack of communication but lower than the *ex ante* probability of peace in the peace talk game.

We conclude that in the mediation game where the distribution function over the set of recommendations is given by (2.50) the best separating equilibrium does not improve on unmediated communication.

2.4.3 Mediation programme mixing for high types

In this section we consider the mediation programme given by

$$F(h,h) = (q_h, p_h, p_h, 1 - q_h - 2p_h), \quad F(h,l) = (q_m, p_m, 0, 1 - q_m - p_m),$$
$$F(l,l) = (q_l, 0, 0, 1 - q_l). \tag{2.69}$$

By symmetry $F(l,h) = (q_m, 0, p_m, 1 - q_m - p_m)$. For this type of programmes the mediator chooses the value of x and probabilities that maximize the probability of peace

$$max_{x,q_l,q_h,q_m,p_l,p_m} \left\{ \frac{1}{4}(q_h + 2p_h) + \frac{1}{2}(q_m + p_m) + \frac{1}{4}q_l \right\}, \tag{2.70}$$

subject to the probability constraints

$$0 \le q_l \le 1,\ 0 \le q_m \le 1,\ 0 \le q_h \le 1,\ 0 \le q_h \le 1, 0 \le p_h \le 1,\ 0 \le p_m \le 1,$$
$$q_m + p_m \ \le 1, q_h + 2p_h \ \le 1. \tag{2.71}$$

The objective function is also subject to IC and IR constraints that follow.

A reported high type may receive from the mediator two types of recommendation. Hence, we consider two *ex post* IR constraints for the high type share. A high type is offered x with probability $p_h + p_m$. After this recommendation he updates his belief about the type of the opponent as $Pr[x \,|m_1 = h] = P[m_2 = h|\ x \wedge m_1 = h] + P[m_2 = l \mid x \wedge m_1 = h]$. We require that war is unprofitable *ex post* for a reported high type who receives a recommendation x. Hence, an *ex post* IR constraint for the high type share states

$$x(p_h + p_m) \geq p_h \pi_h(t_i) + p_m \left(\frac{1}{2} - \frac{\alpha t_i}{8} \right) \tag{2.72}$$

for any $t_i \geq 1/2$. Clearly, inequality (2.72) holds for any p_h and p_m that satisfy the probability constraints because $x > 1/2$ and $\pi_h(t_i) < \frac{1}{2}$ by lemma (2).

A high type is offered $1/2$ with probability $q_h + q_m$. After this recommendation he updates his belief about the type of the opponent as $Pr[1/2 \,|m_1 = h] = P[m_2 = h|\ 1/2 \wedge m_1 = h] + P[m_2 = l \mid 1/2 \wedge m_1 = h]$. Hence, another *ex post* IR constraint for the high type share states

$$\frac{1}{2}(q_h + q_m) \geq q_h \pi_h(t_i) + q_m p \left(\frac{1}{2} - \frac{\alpha t_i}{8} \right)$$

for any $t_i \geq 1/2$. Clearly, this inequality holds for any q_h and q_m which satisfy the probability constraints because $p < 1$ and $\pi_h(t_i) < \frac{1}{2}$ by lemma (2).

A reported low type may receive from the mediator three types of recommendation. Hence, we consider three *ex post* IR constraints for the low type share. A low type is offered $1-x$ with probability $Pr[1-x|m_1 = l] = P[m_2 = h|\ 1-x \wedge m_1 = l] = p_m$. We require war to be unprofitable *ex post* for a low type who receives a recommendation $1-x$ and beliefs that the message sent by the opponent is $m_2 = h$. Hence, another *ex post* IR constraint for the low type share states

$$(1-x) \geq (1-p) \left(\frac{1}{2} - \frac{3\alpha t_i}{8} \right)$$

for any $t_i \leq 1/2$. Hence, an *ex post* IR constraints for the low type share is

$$(1-x) \geq \frac{1}{2}(1-p). \tag{2.73}$$

A reported low type is offered $1/2$ with probability $Pr[1/2|m_1 = l] = P[m_2 = h|\ 1/2 \wedge m_1 = l] + P[m_2 = l|\ 1/2 \wedge m_1 = l] = q_m + q_l$. We require that war is unprofitable *ex post* for a low type who receives a recommendation $1/2$. Hence, another *ex post* IR constraint for the low type share states

$$\frac{1}{2}(q_m+q_l) \geq q_m(1-p)\left(\frac{1}{2}-\frac{3\alpha t_i}{8}\right)+q_l\pi_l(t_i)$$

for any $t_i \leq 1/2$. This IR constraint is obviously satisfied because $\pi_l(t_i) < \frac{1}{2}$ by lemma (2).

Although the recommendation of the mediator can not be enforced, in the construction of the *interim* IC constraints for the mediation game we assume that misreporting is never followed by a failure to comply with recommendation. We will check later in the section if the solution of the program provides some incentive for players to deviate by waging a war after misreporting.

The *interim* IC constraint for the high type requires inequality

$$\frac{1}{2}\left((1-q_h-2p_h)\pi_h(t_i)+q_h\frac{1}{2}+p_hx+p_h(1-x)\right)+ \tag{2.74}$$

$$+\frac{1}{2}\left((1-q_m-p_m)p\left(\frac{1}{2}-\frac{\alpha t_i}{8}\right)+q_m\frac{1}{2}+p_mx\right)\geq$$

$$\geq\frac{1}{2}((1-q_m-p_m)\pi_h(t_i)+q_m\frac{1}{2}+p_m(1-x))+\frac{1}{2}((1-q_l)p(\frac{1}{2}-\frac{\alpha t_i}{8})+q_l\frac{1}{2})$$

to be satisfied for all $t_i \geq 1/2$.

The *interim* IC constraint for the low type requires inequality

$$\frac{1}{2}\left((1-q_m-p_m)(1-p)\left(\frac{1}{2}-\frac{3\alpha t_i}{8}\right)+q_m\frac{1}{2}+p_m(1-x)\right)+ \tag{2.75}$$

$$+\frac{1}{2}\left((1-q_l)\pi_l(t_i)+q_l\frac{1}{2}\right)\geq$$

$$\geq\frac{1}{2}\left((1-q_h-2p_h)(1-p)\left(\frac{1}{2}-\frac{3\alpha t_i}{8}\right)+q_h\frac{1}{2}+p_hx+p_h(1-x)\right)+$$

$$+\frac{1}{2}\left((1-q_m-p_m)\pi_l(t_i)+q_m\frac{1}{2}+p_mx\right)$$

to be satisfied for all $t_i \leq 1/2$.

For the sake of convenience we change the parametrisation of the relaxed program by introducing four variables. We set $s_h \equiv q_h+2p_h$, $s_m \equiv q_m+p_m$ and $s_l \equiv q_l$. Additionally, we denote by

$$b=\frac{1}{2}q_m+xp_m$$

the expected payoff of reported high type in the case of peaceful agreement with a low type. We note that the expected payoff of reported low type in the case of peaceful agreement with a high type is

$$\frac{1}{2}q_m + (1-x)p_m = s_m - b$$

and the expected payoff of reported high type in the case of peaceful agreement with high type is

$$\frac{1}{2}q_h + xp_h + (1-x)p_h = \frac{1}{2}q_h + p_h = \frac{1}{2}s_h.$$

We solve the following relaxed problem:

$$max_{b,s_l,s_m,s_h}\left\{\frac{1}{4}s_h + \frac{1}{2}s_m + \frac{1}{4}s_l\right\}, \tag{2.76}$$

subject to the probability constraints

$$s_l \leq 1,\ 0 \leq s_m \leq 1,\ 0 \leq s_h \leq 1. \tag{2.77}$$

The *interim* IR constraint requires that, after learning its own type but before receiving any recommendation, participation in the mechanism and following all recommendations is at least as profitable for a player as going to war. The *interim* IR constraint for the high type states

$$\frac{1}{2}(s_h\frac{1}{2} + (1-s_h)\pi_h(t_i)) + \frac{1}{2}(b + (1-s_m)p(\frac{1}{2} - \frac{\alpha t_i}{8})) \geq \frac{1}{2}\pi_h(t_i) + \frac{1}{2}p(\frac{1}{2} - \frac{\alpha t_i}{8})$$

for any $t_i \geq 1/2$, which is equivalent to inequality

$$s_h\left(\frac{1}{2} - \pi_h(t_i)\right) + b - s_m p\left(\frac{1}{2} - \frac{\alpha t_i}{8}\right) \geq 0 \tag{2.78}$$

for any $t_i \geq 1/2$.

The *interim* IC constraint for the low type states

$$\frac{1}{2}\left(\frac{1}{2}s_l + (1-s_l)\pi_l(t_i)\right) + \frac{1}{2}\left((s_m - b) + (1-s_m)(1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8})\right) \geq$$

$$\frac{1}{2}(b + (1-s_m)\pi_l(t_i)) + \frac{1}{2}\left(s_h\frac{1}{2} + (1-s_h)(1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8})\right)$$

for any $t_i \leq 1/2$, which is equivalent to inequality

$$s_l\left(\frac{1}{2}-\pi_l(t_i)\right)+s_m-s_m(1-p)\left(\frac{1}{2}-\frac{3\alpha t_i}{8}\right)\geq \tag{2.79}$$

$$2b-s_m\pi_l(t_i)+s_h\left(\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha t_i}{8})\right)$$

for any $t_i \leq 1/2$.

Consider the game G induced by the mediation programme (2.50) where the decision rule satisfies constraints (2.79) and (2.78), probability constraints (2.77), and solves the problem (2.76).

The following proposition is proved in the Appendix.

Proposition 7 *There is an unique best separating equilibrium of the game G induced by the mediation programme* (2.69). *The ex ante probability of peace in this equilibrium does not improve on unmediated communication.*

2.5 Conclusion

We studied several mechanisms that maximise the probability of peaceful settlement between two players contesting a prize of common value. The privately known type of a player describes his strength in a possible conflict. Players hold a common prior belief about the type of their opponent. We depart from the benchmark model by allowing for the possibility of uncertain cost of conflict and continuous types.

We calculated an upper bound for the probability of peaceful split in the case of lack of communication and a simultaneous choice of players whether to agree to a given split proposal. We rank sets of Bayesian Nash equilibria that are achievable in games induced by some information sharing mechanisms. This Pareto ranking of sets is based on the *ex ante* probability of peace in the best separating equilibrium of the induced game.

We study the probability of peace achieved by mechanisms with partial disclosure of private information. We allow for some blurring of the boundaries between partial types. We found that an unmediated cheap talk between players improves chances for peace for any parameter values, even in the case of partial disclosure. In a conflict resolution environment monetary transfers between players are not available as a tool to provide incentives for peaceful settlement. We studied the efficiency of a strategic mediator without enforcement power in the case of a mediation game with the same discrete message space as the unmediated peace talk game. We calculated the probability of peace in the best separating equilibrium of games induced by two incentive compatible mediation programs. One of these mediation programs is characterised by distribution function F defined as follows

$$F(h,h) = (q_h, 0, 0, 1-q_h), \quad F(h,l) = (q_m, p_m, 0, 1-q_m-p_m), \tag{2.80}$$

$$F(l,l) = (q_l, p_l, p_l, 1-q_l-2p_l), \quad F(l,h) = (q_m, 0, p_m, 1-q_m-p_m).$$

In this program the mediator offers with positive probability a higher share of the prise to one of two low types. The other incentive compatible mediation program is characterised by distribution function F defined as follows

$$F(h,h) = (q_h, p_m, p_m, 1-q_h), \quad F(h,l) = (q_m, p_m, 0, 1-q_m-2p_m), \tag{2.81}$$

$$F(l,l) = (q_l, 0, 0, 1-q_l-2p_l), \quad F(l,h) = (q_m, 0, p_m, 1-q_m-2p_m).$$

In this program the mediator offers with positive probability a higher share of the prise to one of two high types. We found that these two mediation programs are not effective in promoting peace. While for some parameter values each of the mediation programs improves chances for peace upon unmediated communication, it worsens the probability of peace compared to direct communication. We found that the *ex ante* probability of peace in the best separating equilibria of the two mediation programs is the same. Mediation programs (2.81) guarantees this maximal *ex ante* probability of peace for a wider range of parameter values than the mediation program (2.80). Hence, any linear combination of the two programs will remain incentive compatible but is not likely to improve chances for peace upon unmediated communication with the same messages space. This result is consistent with findings presented by Fey and Ramsay (2010) in [6]. The authors show that for a broad class of crisis bargaining games with no restriction on the motivation of the mediator any equilibrium outcome that is achievable through mediation is also achievable as an equilibrium outcome of an unmediated cheap talk game.

Our analysis does not preclude the possibility for other mediation programmes with restricted message space to increase the *ex ante* probability of peace in the given setup. In order to determine a mediation program which can improve upon the unmediated communication, a mediation program characterised by distribution function

$$F(h,h) = (q_h, 0, 0, 1-q_h), \quad F(h,l) = (q_m, p_m, p_m, 1-q_m-2p_m), \quad (2.82)$$

$$F(l,l) = (q_l, 0, 0, 1-q_l-2p_l), \quad F(l,h) = (q_m, p_m, p_m, 1-q_m-2p_m),$$

should also be considered. In this mediation program the mediator does not always offer the higher share to the high type facing a lower type. If the only available peaceful mechanism is given by programme (2.82) then it may be optimal for the players to share their information with the mediator in spite of imperfect alignment of interests. However, if players have opportunity to choose a mediator they may avoid mediators that offer this mediation program and might prefer to seek the advice of mediators using programs (2.81) or (2.80).

In our model players are *ex ante* symmetric. In general, types of players can be uniformly distributed in different intervals. It is straightforward to extend the model for the case where the length of the two intervals is the same but players are recognizable by the lower and upper bound for their type. We model this case by letting $T_1 = [0,1]$ and $T_2 = [\beta, 1+\beta]$, where $\beta \in (0,1)$.

A. Meirowitz et al. (2012) show in [13] that mediation might create incentives for players to invest in militarization and increase their war capacity. Hence, the mediator seeking to improve the welfare of the opponents might face a tradeoff between minimization of destruction caused by a possible conflict and minimization of its probability. In this case the decisive information for a mediator is the value of the prize for the winner in a potential current war. This is an example of information that an opponent is willing to share with a strategic mediator but not willing to share with an adversary. The cost $\alpha t_i t_j$ of potential war in our model can be obtained without revealing types t_i and t_j of the opponents. Application of secure function evaluation protocol allows the mediator to evaluate any function $f(t_i, t_j)$, for instance $t_i t_j$, while players keep private their types. In the frame of our model we could evaluate the recommendation of a mediator that has an access to the partial revelation of private information.

References

[1] S. Baliga, T. Sjostrom, *Mechanism Design: Recent Developments,* in L. Blum and S. Durlauf, editors, The New Palgrave Dictionary of Economics (2011)

[2] H. Bester, K. Warneryd, *Conflict and the Social Contract,* Scandinavian Journal of Economics, 108(2) (2006), 231-49.

[3] L. Corchon, M. Dahm, *Fondations for contest success functions,* Economic Theory, 43(1) (2010), 81-98.

[4] J. D. Fearon *Rationalist explanations for war,* International Organization, 49(3) (1995), 379-414.

[5] M. Fey, K. Ramsay, *Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict,* American Journal of Political Science, 55(1) (2011), 149-169.

[6] M. Fey, K. Ramsay, *When is shuttle diplomacy worth the commute? Information sharing through mediation,* World Politics 62.04 (2010): 529-560.

[7] M. Fey, K. Ramsay, *Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types,* Review of Economic Design, 13(3) (2009), 233-250.

[8] M. Garfinkel, S. Skaperdas, *Economics of Conflict: An Overview,* in Todd Sandler and Keith Hartley, editors, Handbook of Defense Economics, Vol. II (2007) Elsevier.

[9] M. Gentzkow, E. Kamenica, *Bayesian persuasion,* No. w15540. National Bureau of Economic Research (2009).

[10] M. Gentzkow, E. Kamenica, *Competition in persuasion,* No. w17436. National Bureau of Economic Research (2011).

[11] J. Horner, M. Morelli, F. Squintani, *Mediation and Peace,* mimeo (2010), Yale University.

[12] M.O. Jackson *Mechanism Theory,* in U. Derigs (ed.) Optimization and Operations Research, Encyclopedia of Life Support Systems (Oxford, U.K.: EOLSS) (2003) http://www.eolss.net

[13] A. Meirowitz, M. Morelli, K. Ramsay, F. Squintani, *Mediation and Strategic Militarization,* mimeo (2012), Columbia University.

[14] K. Warneryd, *Informational Aspects of Conflict,* in M. Garfinkel and S. Skaperdas (eds.), Oxford Handbook of the Economics of Peace and Conflict, Oxford University Press, Chap. 2 (2010).

Chapter 3

Fair allocation of heterogeneous objects with common beliefs about preferences

3.1 Motivation

The study of mechanisms that allocate one object per agent without monetary transfers is motivated by the examples of children's placement in public schools, office or task allocation in institutions, and housing allocation in colleges. In these cases there are no monetary compensations available and there is no possibility for resale.

In many public school admission systems the admissions are decided through a centralized matching market. In the centralized matching market children's legal guardians submit their preference information and the market responds with a matching that optimizes some utility function across all matches. Two commonly desired properties of the matching are efficiency and fairness. Matching is efficient if no other matching gives all children higher utility, under their preferences. One of the most prominent criteria for fairness of matching is lack of envy. Matching is envy-free if each child considers his allocated school at least as valuable as other children's allocation. Admission criteria in many school admission systems allow for a set of children to score equally on the admission test. For example, charter schools in the US grant equal chances to students who reside within the boundaries of a catchment area. Many such schools have coarser criteria which do not create strict ordering of students. For instance, students with siblings at the school are given higher priority than students without such siblings but among students with siblings, all students have equal priority. The problem created by ties is not relevant for each recruitment system. For instance, the notion of catchment area in the school admissions system in England is misleading. According to the study by Burgess et al. (see [3]), "Admission [to public schools in England] is generally not based on within-district vs outside-district criteria, but based more on continuous measures such as degree of proximity and straight line distance". This criteria allows the school admissions system in England to perfectly discriminate between any pair of students (or rather between the wealth of their parents) on the base of the distance of their place of residence to the school, measured in terms of feet. The drawback of this system in which priorities reflect school preferences is a choice restriction for students. Burgess et al. point out in [3] that this problem is particularly severe for the school admissions system in England. Abdulkadiroglu provides in [1] a broad discussion of the school admission problem and various school assignment mechanisms. The author shows that the presence of ties is persistent in the school choice and the question how to break these ties raises some significant design difficulties.

In a symmetric environment all schools have the same capacity and all students tie in priorities at every school. In this case there is no priority ranking of applicants. For the sake of simplicity, assume that there are as many students as different public schools and that each school has a capacity of one student. Students have preferences over schools and do not value money. Indivisible objects are widely allocated by means of lotteries. Consider a random allocation mechanism that solicits students' preferences and represents schools as divisible objects. In this mechanism each student is allocated a lottery that is a probability distribution over deterministic assignments. Budish et al. (2013) generalize the theory of random assignments and determine in [2] when random assignments that satisfy a structure of constraints can be implemented by a lottery over deterministic assignments satisfying a given constraints structure. This theory allows for application of random assignment mechanisms in the case of group-specific quotas in school choice. Efficiency and fairness are two normative criteria that a random assignment is expected to meet. An allocation of lotteries is efficient if no other lottery allocation gives each child a higher expected utility, under their preferences. A lottery allocation is envy-free if each student considers his own lottery at least as valuable as other student's lottery.

In this case no student wishes to swap the lotteries. The mechanism that assigns to each student a lottery with equal probability to be allocated to each school is envy-free but not efficient in the case when some of favourite schools are different. The random serial dictatorship mechanism is appealing because it results in efficient allocation but can create justified envy when some children have similar preferences.

It is not trivial to meet the goal of efficiency because students' preferences are unknown to the admission office. We call an assignment mechanism strategy-proof if a rational student will never find it in his self-interest to falsely report his own preferences. If the mechanism is not strategy-proof, any objectives it might achieve will be with respect to potentially false preferences. We examine in the subsequent paragraphs the existence of efficient, envy-free and strategy-proof random assignment mechanism for different numbers of students that tie in priorities. The existence of such mechanism would help to promote school admissions systems with discrete criteria within the catchment area. Such admission systems offer more options to parents and students and lead to a decrease in social segregation of neighbourhoods.

3.1.1 Two-children ordinal mechanism

Consider a random mechanism that has to allocate two children to public schools where each school can admit only one child. The mechanism may solicit children's ranking of the two most preferred schools. If the first choices are different then each child is assigned his favourite school with probability 1. When the first choices are the same then each child is allocated to this school with probability $1/2$. In this case each child is allocated to his second best school with the remaining probability $1/2$. Truth telling is a dominant strategy for each child participating in this mechanism. The resulting assignment of two lotteries is envy-free and efficient for any von Neumann-Morgenstern utility functions that represent preferences of participants.

3.1.2 Three-children three-school ordinal mechanism

Consider a random mechanism that has to allocate three children to public schools where each school can admit only one child. The mechanism may solicit again children's ordinal preferences which are assumed to be transitive and complete.

The efficiency of an ordinal mechanism is defined with respect to the partial ordering of first order stochastic dominance. Katta and Sethuraman provide an example in [8] showing that any efficient and envy-free ordinal mechanism is not strategy-proof in the full preference domain. For the sake of completeness we present the example given in [8]. Denote the set of schools by (a,b,c) and the set of children by $(1,2,3)$ where the preference profile is given by Table (3.1).

Table 3.1 Preference orderings

Child			
1	{a,b}		c
2	a	b	c
3	a	c	b

The ordinally efficient and envy-free probabilistic assignment for this profile is unique and is the one shown in Table (3.2).

Table 3.2 Probabilistic assignment

Child	a	b	c
1	0	3/4	1/4
2	1/2	1/4	1/4
3	1/2	0	1/2

Assume that child 1 falsely reports preference ordering $\{a,b,c\}$. The only ordinally efficient and envy-free probabilistic assignment for this profile is unique and is the one shown in Table (3.3).

Table 3.3 Probabilistic assignment

Child	a	b	c
1	1/3	1/2	1/6
2	1/3	1/2	1/6
3	1/3	0	2/3

Lottery $(1/3,1/2,1/6)$ assigned to child 1 in the probabilistic assignment in Table (3.3) first order stochastically dominates lottery $(0,3/4,1/4)$ assigned to child 1 in the probabilistic assignment in Table (3.2). Hence, the ordinal mechanism that finds ordinally efficient and envy-free probabilistic assignments is not dominant strategy incentive compatible where we allow for indifference between alternatives. Although the assumption of strict preferences is useful in many applications, in some everyday cases, we do not have, and do not need, strict preferences.

3.1.3 The role of beliefs in the school admission process

In any school admission system the favourite schools are rarely the ones that children actually nominate in their preference lists. Once the admission criteria are set by the local or school authority, parents form well defined expectations about their neighbours' school preferences and about schools that will be oversubscribed. At the time when applications are made, highly desirable schools with small catchment areas or expected oversubscription can be excluded from the choice set. Conversely, some parents might be almost certain that they will be offered a place because there are no much candidates living in the close proximity of the school. The school choice of parents is affected not only by the admission criteria (or the mechanism design) used by schools and local authorities in their area but also by the common beliefs about relative popularity of schools. Parents' awareness of the restricted access to the more popular schools determines to a great extent "truthfulness" of preferences reported to the mechanism. This makes Bayesian implementation particularly relevant. The environment justifies our requirement for the assignment mechanism to be incentive compatible with respect to commonly held beliefs about the joint distribution of peers' preferences.

3.2 The cardinal assignment mechanism

Consider an economy with a set of indivisible objects in which each agent can be allocated exactly one object. Cardinal utilities of agents are private information that are drawn from a commonly known distribution. We study the possibility for a mechanism that solicits cardinal information in agents' preferences to improve over allocation mechanisms that utilize only agents' ordinal preferences. We consider a deterministic mechanism that allows agents to report for each object a limited number of preference intensities. Types of agents are finite and multidimensional. We assume that there is a common knowledge among the agents of a common prior over the preference profiles. A mechanism designer wishes to implement a probabilistic assignment for each realisation of type profiles. We consider the revelation game with incomplete information induced by the mechanism. The strategy proof requirement for the mechanism is relaxed to the requirement of Bayesian incentive compatibility. For any given preference profile and system of beliefs, the mechanism is evaluated by employing the concepts of ex post incentive efficiency and ex post envy-freeness.

3.2.1 Information structure

Let $I = \{1, 2, ..., n\}$ be the set of agents and J the collection of objects, where $|J| = n$. During the ex ante stage each agent privately observes n-dimensional von Neumann-Morgenstern utility index for the set of objects. We assume that individual preference relations can produce multiple minimal and maximal elements for the set of alternatives. We restrict the domain of preferences by fixing the minimal difference in utility level of alternatives. We use the finite utility grid model employed by Dutta, Peters and Sen in [5]. In particular, we allow for three distinct utility levels: 0, some $\eta > 0$, and $1 - \eta > 0$, where $\eta \in (0, 1)$ is chosen by the mechanism designer. The utility of agent i from receiving object j is denoted by $v_i(j)$ and restricted to the domain $\{0, \eta, 1\}$. An admissible von Neumann-Morgenstern utility function for agent i is $v_i : J \to \{0, \eta, 1\}$ where there exists at least one element $j \in J$ such that $v_i(j) = 1$ and at least one element $j' \in J$ such that $v_i(j') = 0$. Denote $v_{ij} = v_i(j)$. Let T_i be the set of admissible vectors of utility valuations $(v_{i1}, ..., v_{in})$, $|T_i| = 3^{n-2}n(n-1)$. The private signal of agent i is a vector $t_i = (v_{i1}, ..., v_{in}) \in T_i$ that will be referred to as agent's type. The type space T_i has a natural partial order that we denote by $\preceq_i$. This order is the product order defined by

$$t_i \preceq_i t_i' \iff t_{ih} \leq t_{ih}', \forall h = 1, ..., n.$$

Define $t_i \prec_i t_i'$ by $t_i \preceq_i t_i'$ and $t_i \neq t_i'$.

Types of agents other than i are denoted by T_{-i} where $T_{-i} \in \mathscr{T}_{-i} \equiv \times_{j \neq i} T_i$. Let T denotes a full type profile where $T \in \mathscr{T} \equiv \times_{i \in I} T_i$.

Information is incomplete. Prior beliefs of agents about the distribution of preferences in the population are represented by cumulative distribution function $F : T \to [0, 1]$. The joint distribution of types F is commonly known. We allow the valuation vectors for agents to be arbitrarily correlated among agents and among objects. Denote by F_i the marginal cumulative distribution of types for each agent i. It is assumed that function $F_i : T_i \to [0,1]$ has positive probability mass function f_i on the support T_i. Denote by $p_{-i}(.|t_i)$ the conditional probability on T_{-i} given t_i. Probability function $p_{-i}(.|t_i)$ will be referred to as the posterior belief of type t_i of agent i about the distribution of others' types.

3.2.2 The allocation mechanism

A feasible assignment matrix $X \in \mathscr{X}$ is a bistochastic matrix where x_{ij} is the objective probability of agent i receiving object j. We denote by x_i the vector of probability shares for agent i in the matrix X.

A pure assignment a is given by a permutation matrix $[a_{ij}]$ such that

$$\forall j \in J \sum_{i=1}^{m} a_{ij} = 1, \; \forall i \in I \sum_{j=1}^{n} a_{ij} = 1 \; and \; a_{ij} \in \{0,1\}. \tag{3.1}$$

The Birkhoff-von Neumann Theorem (1946,1953) shows that any bistochastic matrix is a convex combination of permutation matrices. This theorem implies that, in the case where the number of agents equals the number of objects and agents have unit demand, any feasible assignment is implementable as a lottery over the set of pure assignments.

By revelation principle (see e.g. Myerson [12]), there is no loss of generality in restricting our attention to direct mechanisms. A mechanism designer wishes to allocate the objects for each realization of type profiles. Agents are asked to declare a type and the profile of declared types determines the assignment matrix. The revelation mechanism is denoted by $\Gamma = (\{T_i\}, \phi)$. The objective is represented by function $\phi : \mathscr{T} \to \mathscr{X}$ that takes as input a profile of declared types and returns an assignment matrix in the set $\mathscr{X}$. Let the outcome be $\phi(T) = X$. We adopt a notation where $\phi_i(T)$ is the lottery x_i for agent i selected by the rule ϕ for the type profile T, and $\phi_{ij}(T) = x_{ij}$ denotes the probability share for allocation of object j in this lottery.

The expected utility for type t_i of agent i from assignment X is the inner product $u_i(X,t_i) = x_i \cdot t_i$. We assume that a mechanism designer knows the joint distribution of types F and uses this probability measures to evaluate expected utility of different types of agents.

Denote by ΔT_i the probability distribution on the set T_i. A strategy for agent i in a direct mechanism is a mapping $\psi : T_i \to \Delta T_i$. Truth telling corresponds to the identity matrix $I_{|T_i|}$.

The ex post utility from the rule ϕ for agent i of type t_i when the type profile is T is

$$U^i(\phi(T),t_i) = t_i \cdot \phi_i(T).$$

The interim expected assignment to agent i if he reports type s_i, assuming all other agents truthfully report their type, is defined by

$$\hat{\phi}_i(\phi|s_i) = \sum_{T_{-i}\in\mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_i(T_{-i},s_i).$$

Denote by $U^i(\phi|t_i,s_i)$ the interim expected utility to type t_i of agent i if he reports type s_i, assuming all other agents truthfully report their type, i.e.

$$U^i(\phi|t_i,s_i) = t_i \cdot \hat{\phi}_i(\phi|s_i). \tag{3.2}$$

Denote $U^i(\phi|t_i) = U^i(\phi|t_i,t_i)$ the interim expected utility of agent i from a truthful report.

The Bayesian incentive compatibility requires that reporting the true type be a Bayesian equilibrium of the revelation game induced by Γ.

Definition 5 *A direct mechanism Γ is (interim) Bayesian incentive compatible for the joint distribution of valuation profiles F if for all types t_i of agent i and reports s_i*

$$U^i(\phi|t_i) \geq U^i(\phi|t_i,s_i),\ for\ \forall i. \tag{3.3}$$

We consider the set of incentive compatible mechanisms with the messages space T_i for each agent. We will define below the criteria that can be applied under asymmetric information to select a socially desirable mechanism from this set. *Ex post* refers to the stage at which each agent knows his own preferences and knows other agents' reports.

Definition 6 *A direct mechanism Γ is ex post incentive efficient for the joint distribution of valuation profiles F if it is incentive compatible and there is no incentive compatible direct mechanisms with assignment rule ϕ' such that for all $i \in I$ and all $T \in \mathscr{T}$*

$$t_i \cdot \phi'_i(T) \geq t_i \cdot \phi_i(T)$$

with at least one strict inequality.

Interim refers to the stage at which each agent knows his own preferences and knows how other agents' preferences are distributed only. We consider the situation when the mechanism is selected at the interim stage. The selection criteria in this case is captured by the notion of interim incentive efficiency introduced by Holmstrom and Myerson (1983).

Definition 7 *A direct mechanism* Γ *is interim efficient for the joint distribution of valuation profiles* F *if it is incentive compatible and there is no other mechanism* Γ' *with assignment rule* ϕ' *such that (a)* Γ' *is incentive compatible and (b)* $U^i(\phi'|t_i,t_i) \geq U^i(\phi|t_i,t_i)$ *for all* i*,* t_i *and* $U^i(\phi'|t_k,t_k) > U^i(\phi|t_k,t_k)$ *for some agent* i *of type* t_k*.*

Definition 8 *A direct mechanism* Γ *is ex post envy free for the joint distribution of valuation profiles* F *if for all* i, j *and all type profiles* T

$$t_i \cdot \phi_i(T) \geq t_i \cdot \phi_j(T).$$

Definition 9 *A direct mechanism* Γ *is interim envy free for the joint distribution of valuation profiles* F *if for all* i, j

$$t_i \cdot \hat{\phi}_i(\phi|t_i) \geq t_i \cdot \hat{\phi}_j(\phi|t_i). \tag{3.4}$$

It is easy to show that any ex post envy free mechanism is interim envy free. The set of ex post envy free mechanisms is a proper subset of the set of interim envy free mechanisms. An interim envy free allocations may not be ex post envy free because inequality (3.4) involving linear combinations of utilities $t_i \cdot \phi_i(T_{-i}, s_i)$ does not imply any relationship between components.

We consider two different concepts of monotonicity describing properties of allocation mechanisms that are studied in the following paragraphs.

Definition 10 *A direct mechanism* Γ *is non decreasing in each report for the joint distribution of valuation profiles* F *if for any type* t *of agent* i*, any* $j \in J$*, and any pair of reports* $t_i, t_i' \in T_i$ *such that* $t_{il}' = t_{il}$ *for* $l \neq j$ *and* $t_{ij} > t_{ij}'$ *the assignment rule* ϕ *satisfies condition*

$$\hat{\phi}_{ij}(\phi|t_i) \geq \hat{\phi}_{ij}(\phi|t_i'). \tag{3.5}$$

Definition 11 *A direct mechanism* Γ *is swap monotone on the set* $\mathscr{T}$ *if the allocation probabilities have the same order as the reports of the agents, i.e., for all* $T \in \mathscr{T}$*,* $i,k \in I$*,* $j \in J$ *and* $t_{ij} > t_{kj}$ *the assignment rule* ϕ *satisfies*

$$\phi_{ij}(T) \geq \phi_{kj}(T).$$

The mechanism is said to be individually rational if it is weakly preferred by each agent to the consequences of non-participation. We do not investigate separately the individual rationality of the assignment mechanism. Following the treatment of individual rationality by Crawford in [4], we assume that all assignment mechanisms allow each agent to choose unilaterally to non-participate. Under this assumption, the effect of the unilateral ex ante decision not to participate and the interim decision not to report truthfully is equivalent. Hence, the incentive compatibility property of the mechanism implies the voluntary participation.

3.3 Necessary condition for interim incentive compatibility

We provide necessary condition for Bayesian incentive compatibility of mechanism Γ.

Proposition 8 *A direct mechanism Γ is Bayesian incentive compatible if it is non decreasing in each argument on the set T.*

Proof: To prove necessity, we show that (i) is property of any mechanism Γ satisfying Bayesian incentive compatibility. Let Γ be Bayesian incentive compatible mechanism. Consider any two admissible reports t_i and t_i' for agent i such that $t_{il}' = t_{il}$ for $l \neq j$ and $t_{ij} > t_{ij}'$. Since Γ is incentive compatible, agent i has no incentive to misreport, i.e.

$$t_i \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_i(T_{-i},t_i) \geq t_i \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_i(T_{-i},t_i') \tag{3.6}$$

and

$$t_i' \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i')\ \phi_i(T_{-i},t_i') \geq t_i' \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i')\ \phi_i(T_{-i},t_i). \tag{3.7}$$

Adding terms in each side of inequalities (3.6) and (3.7) yields

$$(t_i - t_i') \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} (p_{-i}(T_{-i}|t_i)\phi_i(T_{-i},t_i) - p_{-i}(T_{-i}|t_i')\phi_i(T_{-i},t_i')) \geq 0. \tag{3.8}$$

Taking into account that $t_i - t'_i = (0, ..., t_{ij} - t'_{ij}, ...0)$ yields

$$(t_{ij} - t'_{ij}) \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} (p_{-i}(T_{-i}|t_i)\phi_{ij}(T_{-i}, t_i) - p_{-i}(T_{-i}|t'_i)\phi_{ij}(T_{-i}, t'_i)) \geq 0. \quad (3.9)$$

As $t_{ij} - t'_{ij} > 0$ inequality (3.9) is equivalent to

$$\sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_{ij}(T_{-i}, t_i) \geq \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t'_i)\ \phi_{ij}(T_{-i}, t'_i). \quad (3.10)$$

Therefore, $\sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_i(T_{-i}, .)$ is a non decreasing function in each argument.

□

Corollary 1.1: If a direct mechanism Γ is Bayesian incentive compatible then for any agent $i \in I$ and any two ordered types $t'_i \preceq_i t_i$

$$U^i(\phi|t'_i) \leq U^i(\phi|t_i). \quad (3.11)$$

Proof: Let Γ be Bayesian incentive compatible mechanism. Then Γ is non decreasing in each argument on the set T and inequality (3.8) is satisfied. Consider any two agents of types $t'_i \preceq_i t_i$. By inequality (3.8) we obtain

$$(t_i - t'_i) \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t'_i)\ \phi_i(T_{-i}, t'_i) \leq (t_i - t'_i) \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_i(T_{-i}, t_i). \quad (3.12)$$

Taking into account inequalities (3.6) and (3.7) we obtain

$$(t_i - t'_i) \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t'_i)\ \phi_i(T_{-i}, t'_i) \leq \quad (3.13)$$

$$\leq U^i(\phi|t_i) - U^i(\phi|t'_i) \leq$$

$$(t_i - t'_i) \cdot \sum_{T_{-i} \in \mathscr{T}_{-i}} p_{-i}(T_{-i}|t_i)\ \phi_i(T_{-i}, t_i).$$

All coordinates of the expected assignment vectors $\hat{\phi}_i(T_{-i}, t'_i)$ and $\hat{\phi}_i(T_{-i}, t_i)$ are non negative. All coordinates of the vector $t_i - t'_i$ have the same sign. Hence, inequality (3.13) implies that the expected utility $U^i(\phi|t_i)$ is non decreasing on $\mathscr{T}$ in each argument of the true type t_i. Inequality (3.13) holds for arbitrary ordered types t_i and t'_i, which establishes (3.11).

□

3.4 The n=3 case

3.4.1 Characterization of interim incentive compatible mechanisms

We provide necessary and sufficient condition for Bayesian incentive compatibility of mechanism Γ for $n = 3$.

Proposition 9 *A direct mechanism Γ is Bayesian incentive compatible iff*

(i) is non decreasing in each argument on the set T;

(ii) for any agent i and any pair of reports $t_i, t_i' \in T_i$, such that $t_{ij} = t_{ij}' = 1$

$$\hat{\phi}_{ij}(T_{-i}, t_i) = \hat{\phi}_{ij}(T_{-i}, t_i');$$

(iii) for any agent i, any pair of reports $t_i, t_i' \in T_i$, and any three different objects $j, k, l \in \{1,2,3\}$, inequalities

$$t_{ij} \geq t_{ik}' \wedge t_{ik} \leq t_{ij}' \wedge t_{il} = t_{il}' = 0,$$

imply

$$\hat{\phi}_{ij}(T_{-i}, t_i) \geq \hat{\phi}_{ik}(T_{-i}, t_i') \quad \textit{and} \quad \hat{\phi}_{ik}(T_{-i}, t_i) \leq \hat{\phi}_{ij}(T_{-i}, t_i').$$

Proof: Sufficiency is proved first. Consider a direct mechanism Γ that satisfies (i) and (ii). For any agent $i \in I$ and any two ordered types $t_i, t_i' \in T_i$ the coordinates of vector $(t_i - t_i')$ have the same sign. Then the expected assignments $\hat{\phi}_i(T_{-i}, t_i)$ and $\hat{\phi}_i(T_{-i}, t_i')$ satisfy

$$(t_i - t_i') \cdot (\hat{\phi}_i(T_{-i}, t_i) - \hat{\phi}_i(T_{-i}, t_i')) \geq 0. \tag{3.14}$$

By adding $(t_i - t_i')\hat{\phi}_i(T_{-i}, t_i')$ to both sides of (3.14) we obtain

$$(t_i - t_i') \cdot (\hat{\phi}_i(T_{-i}, t_i) - \hat{\phi}_i(T_{-i}, t_i') + (t_i - t_i')\hat{\phi}_i(T_{-i}, t_i') \geq (t_i - t_i')\hat{\phi}_i(T_{-i}, t_i'). \tag{3.15}$$

Rearranging terms on both sides of (3.15) yields

$$t_i \cdot (\hat{\phi}_i(T_{-i}, t_i) - \hat{\phi}_i(T_{-i}, t_i')) + t_i' \cdot (\hat{\phi}_i(T_{-i}, t_i') - \hat{\phi}_i(T_{-i}, t_i)) \geq 0. \tag{3.16}$$

Since types t_i and t_i' are arbitrary, both terms in the left side of (3.16) are non negative. Hence,

$$t_i \cdot \hat{\phi}_i(T_{-i}, t_i) \geq t_i \cdot \hat{\phi}_i(T_{-i}, t_i')$$

and

$$t_i' \cdot \hat{\phi}_i(T_{-i}, t_i') \geq t_i' \cdot \hat{\phi}_i(T_{-i}, t_i),$$

which establishes the incentive compatibility of Γ for ordered types.

Let types $t_i, t_i' \in T_i$ be not ordered. Denote by e_k the unit vector in R^3 such that $e_{kk} = 1$. Clearly, $e_k \in T_i$ for any k. There exists $j \in J$ such that $t_{ij} = 1$ and there exists $h \in J$ such that $t_{ih}' = 1$. Let $j = h$. Without loss of generality, let $j = h = 2$ and $t_{i3} = t_{i1} = 0$. Consider $t_i = (x, 1, 0)$ and $t_i' = (0, 1, y)$ where $x, y \in \{\eta, 1\}$. Types $e_2 = (0, 1, 0)$ and t_i' are ordered. Hence,

$$(0,1,y)\hat{\phi}_i(T_{-i}, (0,1,y)) \geq (0,1,y)\hat{\phi}_i(T_{-i}, (0,1,0)).$$

Condition (ii) yields $\hat{\phi}_{i2}(T_{-i}, (0,1,0)) = \hat{\phi}_{i2}(T_{-i}, (x,1,0))$. Condition (i) yields $\hat{\phi}_{i1}(T_{-i}, (0,1,0)) \leq \hat{\phi}_{i1}(T_{-i}, (x,1,0))$. Taking into account that the sum of interim expected shares is equal to one yields

$$\hat{\phi}_{i3}(T_{-i}, (0,1,0)) \geq \hat{\phi}_{i3}(T_{-i}, (x,1,0)).$$

Hence,

$$(0,1,y)\hat{\phi}_i(T_{-i}, (0,1,0)) \geq (0,1,\eta)\hat{\phi}_i(T_{-i}, (x,1,0)).$$

Therefore,

$$(0,1,y)\hat{\phi}_i(T_{-i}, (0,1,y)) \geq (0,1,y)\hat{\phi}_i(T_{-i}, (x,1,0)).$$

Similarly, we notice that

$$(x,1,0)\hat{\phi}_i(T_{-i}, (x,1,0)) \geq (x,1,0)\hat{\phi}_i(T_{-i}, e_2)$$

because types e_2 and t_i are ordered. Condition (ii) yields

$$\hat{\phi}_{i2}(T_{-i}, e_2) = \hat{\phi}_i(T_{-i}, (0,1,y)).$$

By condition (i) $\hat{\phi}_{i3}(T_{-i}, e_2) \leq \hat{\phi}_{i3}(T_{-i}, (0,1,y))$. Hence,

$$\hat{\phi}_{i1}(T_{-i}, e_2) \geq \hat{\phi}_{i1}(T_{-i}, (0,1,y)).$$

Therefore,

$$(x,1,0)\hat{\phi}_i(T_{-i}, e_2) \geq (x,1,0)\hat{\phi}_i(T_{-i}, (0,1,y)).$$

Hence, we obtain

$$(x,1,0)\hat{\phi}_i(T_{-i},(x,1,0)) \geq (x,1,0)\hat{\phi}_i(T_{-i},(0,1,\eta))$$

which establishes that Γ is Bayesian incentive compatible for non ordered types where there exists an object for which both types' value is 1.

Let the non ordered types $t_i, t_i' \in T_i$ be such that $t_{ij} = 1$ and $t'_{ih} = 1$ but $j \neq h$. Without loss of generality, let $t_{i1} = 1$ and $t'_{i3} = 1$. Then $t_i = (1,0,x)$ and $t_i' = (y,0,1)$ where $x, y \in \{0, \eta\}$. Let $x = y$. Condition (iii) yields

$$\hat{\phi}_{i1}(T_{-i},(1,0,x)) = \hat{\phi}_{i1}(T_{-i},(x,0,1)) \quad \textit{and} \quad \hat{\phi}_{i3}(T_{-i},(1,0,x)) = \hat{\phi}_{i3}(T_{-i},(x,0,1)).$$

Hence,

$$(1,0,x)\hat{\phi}_i(T_{-i},(1,0,x)) = (1,0,x)\hat{\phi}_i(T_{-i},(x,0,1))$$

and

$$(x,0,1)\hat{\phi}_i(T_{-i},(x,0,1)) = (x,0,1)\hat{\phi}_i(T_{-i},(1,0,x)).$$

Let $x \neq y$. Without loss of generality, let $x = 0$ and $y = \eta$. Condition (iii) yields

$$\hat{\phi}_{i1}(T_{-i},(1,0,0)) \geq \hat{\phi}_{i1}(T_{-i},(\eta,0,1))$$

and

$$\hat{\phi}_{i3}(T_{-i},(1,0,0)) \leq \hat{\phi}_{i3}(T_{-i},(\eta,0,1)).$$

Hence,

$$(1,0,0)\hat{\phi}_i(T_{-i},(1,0,0)) \geq (1,0,0)\hat{\phi}_i(T_{-i},(\eta,0,1)).$$

Condition (iii) yields also

$$(\eta,0,1)\hat{\phi}_i(T_{-i},(\eta,0,1)) = (\eta,0,1)\hat{\phi}_i(T_{-i},(1,0,\eta)),$$

Taking into account that by condition (ii)

$$\hat{\phi}_{i1}(T_{-i},(1,0,0)) = \hat{\phi}_{i1}(T_{-i},(1,0,\eta))$$

and by condition (i)

$$\hat{\phi}_{i3}(T_{-i},(1,0,0)) \leq \hat{\phi}_{i3}(T_{-i},(1,0,\eta)),$$

we obtain

$$(\eta,0,1)\hat{\phi}_i(T_{-i},(1,0,\eta)) \geq (\eta,0,1)\hat{\phi}_i(T_{-i},(1,0,0)).$$

Therefore,

$$(\eta,0,1)\hat{\phi}_{i1}(T_{-i},(\eta,0,0)) \geq (\eta,0,1)\hat{\phi}_{i1}(T_{-i},(1,0,0))$$

which establishes the Bayesian incentive compatibility of Γ in the case of non ordered types where there does not exist any object for which both types' value is 1.

The necessity of (i) is proved by Proposition (8). To prove necessity of (ii) and (iii), we show that (ii) and (iii) are properties of any Bayesian incentive compatible mechanism Γ. Let Γ be Bayesian incentive compatible mechanism. Without loss of generality, assume that for some agent i and some pair of reports $t_i = (0,1,0)$, $t_i' = (x,1,0)$, where $x \in \{0,\eta\}$, but

$$\hat{\phi}_{ij}(T_{-i},t_i) < \hat{\phi}_{ij}(T_{-i},t_i').$$

Then $t_i\hat{\phi}_{ij}(T_{-i},t_i) < t_i\hat{\phi}_{ij}(T_{-i},t_i')$, which contradicts the incentive compatibility of Γ. This establishes that Γ satisfies (ii). To prove necessity of (iii), consider any pair of reports $t_i,t_i' \in T_i$ such that for some different objects $j,k,l \in \{1,2,3\}$, inequalities

$$t_{ij} \geq t_{ik}' \wedge t_{ik} \leq t_{ij}' \wedge t_{il} = t_{il}' = 0$$

hold. Without loss of generality, let $t_i = (1,0,\eta)$ and $t_i' = (\eta,0,1)$. Since Γ satisfies Bayesian incentive compatibility,

$$\hat{\phi}_{i1}(T_{-i},t_i) + \eta\hat{\phi}_{i3}(T_{-i},t_i) \geq \hat{\phi}_{i1}(T_{-i},t_i') + \eta\hat{\phi}_{i3}(T_{-i},t_i')$$

and

$$\eta\hat{\phi}_{i1}(T_{-i},t_i') + \hat{\phi}_{i3}(T_{-i},t_i') \geq \eta\hat{\phi}_{i1}(T_{-i},t_i) + \hat{\phi}_{i3}(T_{-i},t_i).$$

By rearranging terms in both inequalities, we obtain

$$\hat{\phi}_{i1}(T_{-i},t_i) - \hat{\phi}_{i1}(T_{-i},t_i') \geq \eta(\hat{\phi}_{i3}(T_{-i},t_i') - \hat{\phi}_{i3}(T_{-i},t_i))$$

and

$$\hat{\phi}_{i3}(T_{-i},t_i') - \hat{\phi}_{i3}(T_{-i},t_i) \geq \eta(\hat{\phi}_{i1}(T_{-i},t_i) - \hat{\phi}_{i1}(T_{-i},t_i')).$$

Let $u = \hat{\phi}_{i1}(T_{-i},t_i) - \hat{\phi}_{i1}(T_{-i},t_i')$ and $v = \hat{\phi}_{i3}(T_{-i},t_i') - \hat{\phi}_{i3}(T_{-i},t_i)$. The system of inequalities

$$u \geq \eta v \quad and \quad v \geq \eta u$$

where $0 < \eta < 1$ is satisfied if and only if $u > 0$ and $v > 0$ or $u = v = 0$. Hence, $\hat{\phi}_{i1}(T_{-i},t_i) \geq \hat{\phi}_{i1}(T_{-i},t_i')$ and $\hat{\phi}_{i3}(T_{-i},t_i') \geq \hat{\phi}_{i3}(T_{-i},t_i)$ which establishes (iii).

□

3.4.2 Other necessary conditions for interim incentive compatibility

The following two conditions on the interim utility of types are necessary for the Bayesian incentive compatibility of the mechanism Γ.

Denote by $||t||$ the Euclidean norm of vector t.

Corollary 2.1: A necessary condition for the incentive compatibility of a direct mechanism Γ is the following: for any agent $i \in I$ and any two types $t_i, t_i' \in T_i$, such that $||t_i|| = ||t_i'|| > 1$ and $t_{il} = t_{il}' = 0$ for some $l \in J$, the interim expected utility is the same, i.e.

$$U^i(\phi|t_i') = U^i(\phi|t_i).$$

Proof: Let mechanism Γ be Bayesian incentive compatible. Let types t_i', t_i of agent i be such that $||t_i|| = ||t_i'|| > 1$ and $t_{il} = t_{il}' = 0$. Since types t_i', t_i are feasible, there exists three different objects $j,k,l \in \{1,2,3\}$, such that

$$t_{ij} = t_{ik}' \wedge t_{ik} = t_{ij}' \wedge t_{il} = t_{il}' = 0$$

and $t_{ij} > 0$, $t_{ik} > 0$. By condition (iii) in Proposition 2

$$\hat{\phi}_{ij}(T_{-i},t_i) = \hat{\phi}_{ik}(T_{-i},t_i') \quad and \quad \hat{\phi}_{ik}(T_{-i},t_i) = \hat{\phi}_{ij}(T_{-i},t_i').$$

Therefore, the interim expected utility of types t_i and t_i' satisfy

$$U^i(\phi|t_i) = t_{ij}\hat{\phi}_{ij}(T_{-i},t_i) + t_{ik}\hat{\phi}_{ik}(T_{-i},t_i) = t_{ik}'\hat{\phi}_{ik}(T_{-i},t_i') + t_{ij}'\hat{\phi}_{ij}(T_{-i},t_i') = U^i(\phi|t_i')$$

which completes the proof. □

The next constraint represents the interim utility of a type in terms of the expected value of that type's allocation, given that all agents report truthfully. In the case of one dimensional continuous types model where the utility function is continuously differentiable in types, the continuous types form of this constraint is a part of the two conditions that completely characterize the incentive compatible mechanisms (see e.g. [10] by Ledyard and Palfrey (2007)). In our model with discrete, three dimensional types, this constraint is satisfied "locally" for any subset of ordered types.

Corollary 2.2: A necessary condition for Bayesian incentive compatibility of a direct mechanism Γ is the following: for any agent $i \in I$ and any type $t_i \in T_i$

(i)

$$U^i(\phi|t_i) = U^i(\phi|t_i^0) + (t_i - t_i^0)\hat{\phi}_i(T_{-i}, t_i)$$

where $||t^0|| = 1$ and $t^0 \preceq_i t_i$.

Proof: Let Γ be a Bayesian incentive compatible mechanism. Consider types $t_i, t_i^0 \in T_i$ of agent i such that $||t^0|| = 1$ and $t^0 \preceq_i t_i$. Let $t_{ij} = t_{ij}^0 = 1$. Since Γ is Bayesian incentive compatible, condition (ii) in Proposition 2 is satisfied. Hence,

$$\hat{\phi}_{ij}(T_{-i}, t_i) = \hat{\phi}_{ij}(T_{-i}, t_i').$$

Without loss of generality, let $j = 1$, i.e. $t_i^0 = (1,0,0)$, and let $t_{i3} = t_{i3}' = 0$. Taking into account that $\hat{\phi}_{i1}(T_{-i}, t_i) = \hat{\phi}_{i1}(T_{-i}, t_i^0)$ and $t_i - t_i^0 = (0, t_{i2}, 0)$, we obtain

$$t_{i1}\hat{\phi}_{i1}(T_{-i}, t_i) + t_{i2}\hat{\phi}_{i2}(T_{-i}, t_i) = t_{i1}^0\hat{\phi}_{i1}(T_{-i}, t_i^0) + t_{i2}\hat{\phi}_{i2}(T_{-i}, t_i)$$

which establishes (i). $\square$

3.5 Incentive efficient mechanisms

In the context of general collective choice problem the theorem by Myerson (1983) (see [13]) characterizes the set of interim efficient mechanisms as solutions to a constrained welfare optimisation problem. In this context the constraint guarantees that the expected utility of any agent is a non decreasing function of agent's type.

Denote by λ the set of type dependent interim utility weights $\{\lambda_i : T_i \to R_+\}_{i \in I}$ and by $\lambda_i^0 = \sum_{t_i \in T_i} f_i(t_i)\lambda_i(t_i)$ the agent i's ex ante welfare weight relative to other agents. We state the following result without proof.

Theorem: A direct mechanism Γ is interim efficient if and only if there exists a set of utility weights λ, where $\lambda_i^0 > 0$ for some i, such that Γ maximises

$$\sum_{i\in I}\sum_{t_i\in T_i} \lambda_i(t_i) f_i(t_i) U^i(\phi, t_i)$$

subject to Γ is Bayesian incentive compatible.

Conditions under which the solution of the relaxed optimization problem satisfies the missing constraint is known as the regular case. In some environments sufficient conditions for regularity can be established. For example, Ledyard and Palfrey (2007) provide in [10]) sufficient conditions for regularity for independent linear environments.

Consider the solution of the relaxed optimisation problem

$$max_{\lambda} \sum_{i\in I}\sum_{t_i\in T_i} \lambda_i(t_i) f_i(t_i) U^i(\phi, t_i). \tag{3.17}$$

Since types are independent, the relaxed optimisation problem can be solved in two steps (a linear utility example is provided by Wilson (1992) in [16]). In the first step we optimise group's sum of interim utilities for three groups of types. Each group consists of ordered types comparable with one of the feasible types with norm 1: $(1,0,0),(0,1,0),(0,0,1)$. Clearly, these groups correspond to overlapping sets with common types $(1,1,0),(0,1,1),(1,0,1)$. The weight of a group's sum of expected utilities in the total sum is obtained by assigning uniform constant welfare weights, e.g. 1, for each member of the group. In the second step we optimise the sum of interim utilities within each group.

The following claim is proved in Appendix B.

Proposition 10 *Let Γ be Bayesian incentive compatible. Let λ be the optimal set of type dependent interim utility weights for the optimization problem* (3.17). *Then for any agent i and any types $t_i'' \preceq_i t_i \preceq_i t_i'$, the optimal weights' products $\{w_i(t_i)\}_{t_i\in T_i} = \{\lambda_i(t_i) f_i(t_i)\}_{t_i\in T_i}$ satisfy*

$$w_i(t_i'') \le w_i(t_i) \le w_i(t_i').$$

Holmstrom and Myerson (1983) show that interim efficient mechanisms are also ex post incentive efficient. In the next paragraphs we consider ex post incentive efficient mechanisms.

3.6 Ex post incentive efficient and envy-free mechanisms

We consider the existence of ex post incentive efficient and ex post envy-free direct mechanism Γ. We limit our consideration to the class of deterministic mechanisms. The conditions determining ex post efficient and ex post envy free assignment for some type profile are independent of the probability distribution of types. In contrast, the incentive compatibility constraint is defined with respect to some particular distribution of type profiles F. The set of ex post efficient and envy free assignment rules is not finite although the set of type profiles is finite. In this paragraph we provide examples of type profiles for which there are infinitely many ex post efficient and envy free assignments. There is a lower and upper bound for the utility of each type in any of these assignments. Hence, we expect that for any assignment rule, that is deterministic for each type profile, there exist some distribution of type profiles for which the rule is not incentive compatible.

Recall that any ex post envy-free mechanism is interim envy-free. Let us consider two agents that have the same preferences. Since our mechanism does not create any justified envy, such agents should receive assignments that they deem equivalent. Clearly, an interim envy-free assignment rule satisfies a weak form of anonymity: an equal treatment of equals. It means that ties are resolved by providing the same assignment for any two agents that submit the same report. For every type profile T denote by T^{ij} the type profile obtained from T by exchanging agent i's and agent j's reports.

Definition 12 *A direct mechanism Γ is (weakly) anonymous if for any agents i and j and any type profile $T \in \mathscr{T}$, $\phi_i(T^{ij}) = \phi_j(T)$.*

We note that this is a weak version of anonymity because we do not permute all agents and we do not require that the assignments of other agents remain the same.

3.6.1 Sets of ex post efficient and envy-free assignments

On the full preference domain the ex post ordinal efficient and envy-free assignment is unique for each preference profile (see [8] by Katta and Sethuraman (2006)). The following examples show that for some cardinal type profiles from the set $\mathscr{T}$ there exist infinitely many ex post efficient and envy free assignments.

Katta and Sethuraman show in [8] that probabilistic serial mechanism is not strategy-proof on the full preference domain by using an example of two particular preference profiles. The ordinally efficient and envy-free allocations for these profiles are uniquely determined. First, we calculate the sets of ex post efficient and envy free assignments for two cardinal type profiles that correspond to the preference profiles considered by Katta and Sethuraman. We show that the type of agent for which the ordinal ex post efficient and envy free mechanism is not strategy-proof, has no incentive to misreport in an ex post efficient and envy free mechanism Γ.

For the sake of convenience in the subsequent paragraphs we denote the set of objects by (a,b,c). Let the type profile T be given by Table (3.4).

Table 3.4 Valuations

Agent	a	b	c
1	1	1	0
2	1	η	0
3	1	0	η

Ex post efficiency implies that at least one of the shares x_{1b} and x_{2b} is positive. Then if $x_{3b} > 0$ there exists a Pareto improving trade between agent 3 and agents 1 or 2. Hence, the efficiency of assignment X implies $x_{3b} = 0$.

If $x_{1a} > 0$ then $x_{1b} < 1$. Hence, $x_{2b} > 0$.Then there exists a Pareto improving trade between agents 1 and 2: agent 1 can give up part of his share x_{1a} to agent 2 in exchange to equal part of his share x_{2b}. Hence, efficiency implies $x_{1a} = 0$.

Then the structure of ex post efficient assignment X is

$$X = \begin{bmatrix} 0 & x_{1b} & 1-x_{1b} \\ x_{2a} & 1-x_{1b} & x_{1b}-x_{2a} \\ 1-x_{2a} & 0 & x_{2a} \end{bmatrix}$$

where x_{1b} and x_{2a} satisfy the following envy free conditions.

Agent 2 will not envy agent 1 if

$$\eta x_{1b} \leq x_{2a} + \eta(1-x_{1b}) \iff 2\eta x_{1b} \leq x_{2a} + \eta.$$

Similarly, agent 1 will not envy agent 2 if

$$x_{1b} \geq x_{2a} + (1 - x_{1b}) \iff 2x_{1b} \geq 1 + x_{2a}.$$

Hence, $1/2 < x_{1b} \leq 1$. The following condition guarantees that agent 3 does not envy agent 2

$$1 - x_{2a} + \eta x_{2a} \geq x_{2a} + \eta(x_{1b} - x_{2a}) \iff 1 - \eta x_{1b} \geq 2(1-\eta)x_{2a}.$$

Agent 2 will not envy agent 3 if

$$x_{2a} + \eta(1 - x_{1b}) \geq 1 - x_{2a} \iff 2x_{2a} \geq 1 - \eta + \eta x_{1b}.$$

Agent 1 will not envy agent 3 if

$$1 \leq x_{1b} + x_{2a}.$$

The subset of assignments that correspond to ex post efficient and envy free assignments for the type profile given in Table (3.4) is not convex for each value of the grid parameter η. This subset is shown on Figure (3.1).

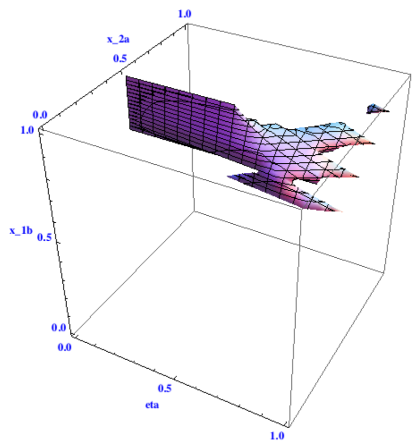

Fig. 3.1 The set of ex post efficient and envy free assignments for the type profile given in Table (3.4)

For example, it is easy to check that this subset is not convex for $\eta = 9/10$. For $\eta = 1/2$ the set of ex post efficient and envy free assignments for the type profile given in Table (3.4) is convex. This set has three extreme points given by pairs of values $(13/28, 6/7)$, $(4/7, 6/7)$, and $(1/2, 1)$ for x_{2a} and x_{1b}:

$$X_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix},$$

$$X_2 = \begin{bmatrix} 0 & 6/7 & 1/7 \\ 13/28 & 4/28 & 11/28 \\ 15/28 & 0 & 13/28 \end{bmatrix},$$

$$X_3 = \begin{bmatrix} 0 & 6/7 & 1/7 \\ 4/7 & 1/7 & 2/7 \\ 3/7 & 0 & 4/7 \end{bmatrix}.$$

Next, we calculate the assignment for agent 1 with report $t_1' = (1, \eta, 0)$ for the type profile T given in Table (3.5).

Table 3.5 Valuations

Agent	a	b	c
1	1	η	0
2	1	η	0
3	1	0	η

Assume that $x_{3b} > 0$. Then $x_{3c} < 1$. It implies that at least one of inequalities $x_{2c} > 0$ and $x_{1c} > 0$ is satisfied. Hence, there exists a Pareto improving trade between agent 3 and agents 2 or 1. Therefore, a necessary condition for ex post efficiency of the assignment for type profile in Table (3.5) is $x_{3b} = 0$. Agents 1 and 2 do not envy each other if

$$x_{1a} + \eta x_{1b} = x_{2a} + \eta x_{2b}.$$

It is easy to show that in an ex post efficient and envy free assignment agents 1 and 2 with identical reports may receive different assignments vectors $(u, v, 1 - u - v)$ and $(u + \eta(2v - 1), 1 - v, v + \eta - u - 2v\eta)$, where $x_{1a} = u$ and $x_{1b} = v$, with equal probability. However, by assumption mechanism Γ satisfies the equal treatment of equals property. Hence, we set $x_{1a} = x_{2a}$ and $x_{1b} = x_{2b}$.

Therefore, the structure of assignment matrix X is the following:

$$X = \begin{bmatrix} x_{1a} & 1/2 & 1/2 - x_{1a} \\ x_{1a} & 1/2 & 1/2 - x_{1a} \\ 1 - 2x_{1a} & 0 & 2x_{1a} \end{bmatrix}. \tag{3.18}$$

Agents 1 and 3 do not envy each other if and only if

$$x_{1a} + \eta \frac{1}{2} \geq 1 - 2x_{1a}$$

and

$$1 - 2x_{1a} + 2\eta x_{1a} \geq x_{1a}.$$

Hence, the assignment matrix (3.18) corresponds to an efficient and envy free assignment if

$$\frac{2-\eta}{6} \leq x_{1a} \leq \frac{1}{3-2\eta}.$$

Therefore, the set of efficient and envy free assignments for type profile (3.5) is convex for any value of η. In particular, for $\eta = 1/2$ the assignment (3.18) is efficient and envy free for any $x_{1a} \in [\frac{1}{4}, \frac{1}{2}]$. The two extreme points of the set of efficient and envy free assignments for $\eta = 1/2$ are

$$X_4 = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{bmatrix}, X_5 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Consider a joint distribution of types where the probability of type profile in Table (3.5) is high. In order to satisfy the condition in Corollary 2.1, an incentive compatible mechanism may have to select an efficient and envy free assignment in which the ex post expected utility of each agent is the same. We obtain such an assignment for $x_{1a} = \frac{2-n}{6-4n}$. For $\eta = 1/2$, the ex post expected utility of type $t_1 = (1,1,0)$ of agent 1 in this assignment is $7/8$. Hence, if the mechanism Γ selects assignment X_1 for the type profile (3.5) then the type t_1 of agent 1 has no incentive to misreport even when the probability of type profile $T_{-1} = ((1,\eta,0),(1,0,\eta))$ is high. Clearly, type $t_1' = (1,\eta,0)$ of agent 1 has no incentive to misreport for any choice of efficient and envy free assignments for the two type profiles.

An example of a type profile with unique ex post efficient and envy free assignment is given in Table (3.7).

Table 3.6 Valuations

Agent	a	b	c
1	1	1	0
2	1	η	0
3	1	η	0

Since mechanism Γ satisfies the equal treatment of equals criterion, in order to calculate the assignment we set $x_{2a} = x_{3a}$, $x_{2b} = x_{3b}$, and $x_{2c} = x_{3c}$. The sufficient conditions for no envy assignment are

$$x_{1a} + x_{1b} \geq x_{2a} + x_{2b} \tag{3.19}$$

and

$$x_{2a} + \eta x_{2b} \geq x_{1a} + \eta x_{1b}.$$

Inequalities (3.19) imply

$$\eta (x_{2b} - x_{1b}) \geq x_{1a} - x_{2a} \geq x_{2b} - x_{1b}.$$

Hence,

$$x_{2b} \leq x_{1b}, \quad x_{1a} \leq x_{2a}$$

and

$$x_{1b} - x_{2b} \leq \frac{x_{2a} - x_{1a}}{\eta}. \tag{3.20}$$

Condition (3.20) binds for an ex post efficient assignment. Otherwise by increasing x_{1b} and reducing x_{2b} we could increase the utility of agent 2 without reducing the utility of agent 1. Hence, $x_{1b} - x_{2b} = \frac{x_{2a} - x_{1a}}{\eta}$. If $x_{1a} > 0$ there is a Pareto improving trade between agents 1 and 2. Therefore, $x_{1a} = 0$ in an ex post efficient assignment. Hence,

$$x_{1b} - x_{2b} = \frac{x_{2a}}{\eta}, \quad \textit{and} \quad x_{1a} = 0. \tag{3.21}$$

Since

$$x_{2a} = x_{3a} \quad \textit{and} \quad x_{2b} = x_{3b},$$

we obtain

$$x_{1b} - x_{2b} = \frac{x_{2a}}{\eta}, \quad 2x_{2a} = 1, \quad x_{1b} + 2x_{2b} = 1.$$

Therefore,

$$x_{2b} = \frac{1}{3} - \frac{1}{6\eta}, \quad x_{1b} = \frac{1}{3} + \frac{1}{3\eta}.$$

Hence, the only ex post efficient and envy free assignment for the type profile T given in Table (3.7) is

$$X_6 = \begin{bmatrix} 0 & \frac{1}{3} + \frac{1}{3\eta} & \frac{2}{3} - \frac{1}{3\eta} \\ 1/2 & \frac{1}{3} - \frac{1}{6\eta} & \frac{1}{6} + \frac{1}{6\eta} \\ 1/2 & \frac{1}{3} - \frac{1}{6\eta} & \frac{1}{6} + \frac{1}{6\eta} \end{bmatrix}.$$

We obtain that for any value of the grid parameter η the ex post efficient and envy free assignment for the type profile given in Table (3.4) is unique.

3.6.2 On the impossibility to achieve fairness in the ex post incentive efficient mechanism

For some joint distributions of types the ex post incentive efficiency and no envy conditions are incompatible.

Proposition 11 *Let deterministic mechanism Γ be anonymous, ex post efficient, envy free, and satisfies equal treatment of equals property. There exist some joint distributions of types such that Γ is not incentive compatible for any value of the grid parameter η.*

Proof: The proof is by example. The marginal distribution f_i of types of agent $i \in I$ will be called object-symmetric if it satisfies conditions

$$f_i((1,0,0)) = f_i((0,1,0)) = f_i((0,0,1)) = f_i^1, \tag{3.22}$$

$$f_i((1,1,0)) = f_i((0,1,1)) = f_i((1,0,1)) = f_i^2,$$

$$f_i((1,\eta,0)) = f_i((\eta,1,0)) = f_i((0,1,\eta)) =$$

$$= f_i((0,\eta,1)) = f_i((1,0,\eta)) = f_i((\eta,0,1)) = f_i^3.$$

We call the joint distribution F object-symmetric if marginal distribution of types f_i is object-symmetric for every $i \in I$. Consider the following joint distribution F in which marginal distributions of types of agents are different. Marginal distributions of types of agent 1 and 3 are identical and object-symmetric. Denote $f^1 = f_1^1 = f_3^1$, $f^2 = f_1^2 = f_3^2$, $f^3 = f_1^3 = f_3^3$. The distribution f_2 of types of agent 2 is not object-symmetric. Probability function f_2 satisfies the last two conditions in (3.22), but not the first one. In particular, let $f_2^2 = f^2$ and $f_2^3 = f^3$. Let $f_2((1,0,0)) = x$, $f_2((0,1,0)) = y$, $f_2((0,0,1)) = z$, where $x, y, z \in [0,1]$, $x \neq y$, $x \neq z$, $y \neq z$, and $x + y + z = 3f^1$.

Consider types $t_2 = (1,\eta,0)$ and $t_2' = (\eta,1,0)$ of agent 2 where $||t_2|| = ||t_2'|| > 1$ and $t_{23} = t_{23}' = 0$. Since Γ is incentive compatible, by Corollary 2.1 the interim expected utilities of types t_2 and t_2' satisfy

$$U^2(\phi|(1,\eta,0)) = U^2(\phi|(\eta,1,0)).$$

By the choice of distribution F, for any $t_1, t_3 \in T_i$ the probability of type profiles $T = t_1 \times (1,\eta,0) \times t_3$ and $T' = t_1 \times (\eta,1,0) \times t_3$, satisfy condition $p(T) = p(T')$. Hence,

$$(1,\eta,0) \sum_{T_{-2} \in \mathscr{T}_{-2}} p_{-2}(T_{-2}|(1,\eta,0))\ \phi_2(T_{-2},(1,\eta,0)) = \tag{3.23}$$

$$= (\eta,1,0) \sum_{T_{-2} \in \mathscr{T}_{-2}} p_{-2}(T_{-2}|(\eta,1,0))\ \phi_2(T_{-2},(\eta,1,0))$$

where $p_{-2}(T_{-2}|(1,\eta,0)) = p_{-2}(T_{-2}|(\eta,1,0))$ for any T_{-2}. Since mechanism Γ is deterministic, all assignments $\phi_2(T_{-2},(1,\eta,0))$ and $\phi_2(T_{-2},(\eta,1,0))$ in equality (3.23) are uniquely determined. These assignments are fixed by Γ not only for type profiles for which the ex post efficient and envy free assignment is unique, but also for type profiles like the type profile given in Table (3.5), for which there exist infinitely many ex post efficient and envy free assignments.

Consider types $t_1 = (1,\eta,0)$ and $t_1' = (\eta,1,0)$ of agent 1. Since Γ is anonymous and satisfies equal treatment of equals property, for any $t_1, t_3 \in T_i$

$$\phi_1((1,\eta,0) \times t_1 \times t_3) = \phi_2(t_1 \times (1,\eta,0) \times t_3)$$

and

$$\phi_1((\eta,1,0) \times t_1 \times t_3) = \phi_2(t_1 \times (\eta,1,0) \times t_3).$$

However, for type profiles $T = (1,\eta,0) \times t_2 \times t_3$ where $t_2 \in \{(1,0,0),(0,1,0),(0,0,1)\}$

$$p_{-1}(T_{-1}|(1,\eta,0)) \neq p_{-2}(T_{-2}|(1,\eta,0))$$

and for type profiles $T = (\eta, 1, 0) \times t_2 \times t_3$ where $t_2 \in \{(1,0,0),(0,1,0),(0,0,1)\}$

$$p_{-1}(T_{-1}|(\eta,1,0)) \neq p_{-2}(T_{-2}|(\eta,1,0)).$$

Assignments for types $t_1 = (1, \eta, 0)$ and $t_1' = (\eta, 1, 0)$ in ex post efficient and envy free mechanism are different for some type profiles where

$$t_2 \in \{(1,0,0),(0,1,0),(0,0,1)\}.$$

For example, for the type profile T_{-1} given in Tables (3.7) and (3.8)

Table 3.7 Valuations

	Objects		
Agent	a	b	c
1	1	η	0
2	1	0	0
3	0	0	1

Table 3.8 Valuations

	Objects		
Agent	a	b	c
1	η	1	0
2	1	0	0
3	0	0	1

the ex post efficient and envy free assignment for type $(1, \eta, 0)$ of agent 1 belongs to the set

$$\{(x_{1a}, x_{1b}, x_{1c}) : \frac{1-\eta}{2-\eta} \leq x_{1a} \leq \frac{1}{2} \wedge x_{1b} = 1 - x_{1a} \wedge x_{1c} = 0\}$$

while the ex post efficient and envy free assignment for type $(\eta, 1, 0)$ is $(0, 1, 0)$. Therefore, the ex post utility of type t_1 is $t_1 \cdot \phi_1(T_{-1}, (1, \eta, 0)) \leq \frac{1}{2} + \frac{1}{2}\eta$ while the ex post utility of type t_1' is $t_1' \cdot \phi_1(T_{-1}, (\eta, 1, 0)) = 1$. Hence, for any $\eta < 1$

$$t_1' \cdot \phi_1(T_{-1}, (\eta, 1, 0)) > t_1 \cdot \phi_1(T_{-1}, (1, \eta, 0)).$$

Taking into account equality (3.23), we conclude that there exist $x, y, z \in [0,1]$ such that $x+y+z = 3f^1$ and

$$(1,\eta,0) \sum_{T_{-1} \in \mathscr{T}_{-1}} p_{-1}(T_{-1}|(1,\eta,0))\ \phi_1(T_{-1},(1,\eta,0)) \neq$$

$$(\eta,1,0) \sum_{T_{-1} \in \mathscr{T}_{-1}} p_{-1}(T_{-1}|(\eta,1,0))\ \phi_1(T_{-1},(\eta,1,0)).$$

Therefore,

$$U^1(\phi|(1,\eta,0)) \neq U^1(\phi|(\eta,1,0)),$$

which violates the necessary condition for the incentive compatibility of Γ established by Corollary 2.1. Therefore, for any value of parameter η, mechanism Γ is not incentive compatible for the joint distribution F. □

The example suggests that the mechanism Γ may be ex post incentive efficient and envy free when the joint distribution F is both agent-symmetric and object-symmetric.

3.7 Conclusion

Our motivation for this project was to understand the ability of cardinal mechanism without monetary transfers to provide efficient and envy-free assignments of heterogeneous objects to agents with private valuations and common prior over the preference profiles. We use a simple assignment mechanism that restricts the message space by allowing agents to report a limited number of preference intensities and indifferences between alternatives. The mechanism assigns a probability distribution over the set of feasible allocations for each preference profile. We provided necessary conditions for incentive compatibility of these mechanisms for any finite number of agents and we characterized the set of incentive compatible assignment rules for the case of three agents. We provided some necessary conditions for regularity of the optimization problem related to the interim efficient mechanisms. We showed that the ex post envy-freeness property may be incompatible with ex post incentive efficiency of the mechanism. We provided an example of a common belief about the distribution of type profiles where the set of ex post incentive efficient and ex post envy free allocations is empty.

The existence of a common prior that describes the joint distribution of agents' types and induces interim beliefs is assumed in much of the applied mechanism design. This assumption eliminates the possibility that agents may assign zero probability to the true state of the world. The reliance of the Bayesian mechanism design on the common prior assumption has been criticized as unrealistic in the literature since the famous paper by Wilson (1995). The critique is focused on environments where the mechanism designer has no information about the beliefs of agents, while for the school allocation problem this is not quite relevant.

Nevertheless, the common prior assumption is not without loss of generality for the designer's problem. Recent literature on mechanism design suggests different methods for overcoming the dependence of the social choice implementation on the common prior assumption. One method is to replace the Bayesian incentive compatibility with a stronger solution concept that is not sensitive to agents' beliefs. For example, the designer may choose a dominant strategy mechanism which is optimal with respect to his subjective beliefs. Another method is to relax the common prior assumption by allowing for the possibility that agent's beliefs do not uniquely determine their preferences. Agent's private information (type) may be informative about or correlated with other agents' types in two senses. One is that agent's preferences may be correlated with agent's own belief, the other is that the prior that describes the joint distribution of agents' types may show correlation among different types. A further study is required to answer the question of which relaxing assumption about the beliefs should be made in the context of the school allocation problem.

It has been well understood since the work by Moore and Repullo (1988) that the set of implementable social choice rules can be dramatically expanded by the use of extensive game forms. One direction for further research can be analysis of decentralized assignment mechanisms that induce extensive game forms and might solve the designer's problem of achieving an efficient and fair allocation.

References

[1] A. Abdulkadiroglu *School Choice,* In Zvika Neeman, Muriel Niederle, Alvin E. Roth, and Nir Vulkan editors, *Oxford Handbook of Market Design*, Oxford University Press (2013).

[2] E. Budish, Y. Che, F. Kojima, P. Milgrom. *Designing Random Allocation Mechanisms: Theory and Applications,* The American Economic Review, Volume 103, Number 2 (2013) : 585-623.

[3] S. Burgess,E. Greaves,A. Vignoles,D.Wilson *What parents want: school preferences and school choice,* The Economic Journal, DOI: 10.1111/ecoj.12153.

[4] V.P. Crawford *Efficient and durable decision rules: A reformulation ,* Econometrica 53 (1985): 817-836.

[5] B. Dutta, H. Peters, A. Sen *Strategy-proof Cardinal Decision Schemes,* Social Choice and Welfare 28 (2007): 163-179.

[6] T. Gajdos, J.M. Tallon *Fairness under Uncertainty,* Economics Bulletin, 4 (2002): 1-7.

[7] G. Hardy, J. Littlewood, G. Polya, *Inequalities,* Cambridge University Press, 1952.

[8] A. Katta, J. Sethuraman *A solution to the random assignment problem on the full preference domain* Journal of Economic Theory 131 (2006): 231-250.

[9] J. Ledyard, T. Palfrey *A characterization of interim efficiency with public goods,* Econometrica 67 (1999): 435-848.

[10] J. Ledyard, T. Palfrey *A general characterization of interim efficient mechanisms for independent linear environments,* Journal of Economic Theory 133 (2007): 441-466.

[11] J. Moore, R. Repullo *Subgame perfect implementation,* Econometrica: Journal of the Econometric Society (1988): 1191-1220.

[12] R. B. Myerson *Incentive Compatibility and the Bargaining Problem* Econometrica 47 (1979): 61-73.

[13] R. B. Myerson *Bayesian equilibrium and incentive-compatibility: An introduction,* Social goals and social organization: Essays in memory of Elisha Pazner (1985): 229-260.
$https://129.105.97.149/research/math/papers/548.pdf$

[14] T. Sonmez, M. U. Unver, *Course bidding at Business Schools,* International Economic Review 51(2010): 99-123.

[15] W. Thompson, *Fair Allocation Rules* Chapter 21 in Handbook of Social Choice and Welfare, vol. 2 (2011): 393-506.

[16] R. Wilson, *Design of Efficient Trading Procedures,* In D. Friedman, J. Geankoplos, D. Lane, J. Rust, *The Double Auction Market* , SFI

Chapter 4

Appendix A

Proof of Lemma 1

An offer of x does not affect the belief of player i about the unknown type t_j and player's expected war payoff $\pi(t_i)$. By Lemma (1), function $\pi(t_i)$ is strictly increasing in t_i on $[0,1]$. Hence, it has an inverse function, which we denote by $\phi(.)$. The inverse function satisfies $\pi(\phi(y)) = y$ for any $y \in [1-p, p(1-\frac{\alpha}{2})]$. The value of $\phi(y)$ is the type of player who has an expected war payoff y. As a monotone real-valued function, $\pi(t_i)$ is differentiable almost everywhere. Similarly, the inverse function $\phi(.)$ is also differentiable almost everywhere. The inverse functions is monotone and increasing in $[1-p, p(1-\frac{\alpha}{2})]$, so we keep the direction of inequality as

$$x \geq \pi(t_i) \Leftrightarrow \phi(x) \geq t_i.$$

Therefore,

$$P(x \geq \pi(t_i)) = P(\phi(x) \geq t_i) = \phi(x).$$

First, we calculate the probability that an offer of $x \in [0,1]$ is accepted by player of type t. Inequality

$$x \geq \pi(t)$$

is equivalent to

$$(1-p-x) + (-1-\frac{\alpha}{2}+2p+\frac{\alpha p}{2})t + (\frac{\alpha}{2}-\alpha p)t^3 \leq 0.$$

Then we solve

$$(2-2p-2x) + (-2-\alpha+4p+\alpha p)t + \alpha(1-2p)t^3 \leq 0. \qquad (4.1)$$

Recall that $1-2p<0$. Dividing both sides by $\alpha(1-2p)$ we obtain

$$t^3+\frac{-2-\alpha+4p+\alpha p}{\alpha(1-2p)}t+\frac{(2-2p-2x)}{\alpha(1-2p)}\geq 0. \tag{4.2}$$

Equation

$$t^3+\frac{-2-\alpha+4p+\alpha p}{\alpha(1-2p)}t+\frac{(2-2p-2x)}{\alpha(1-2p)}=0 \tag{4.3}$$

has three distinct real roots. Hence, although all roots are real, we require complex numbers to express them in radicals. That is why we express the root which belongs to the interval $[0,1]$ in terms of the cos and arccos functions. For simplicity of notation we denote

$$r=\frac{-2-\alpha+4p+\alpha p}{\alpha(1-2p)} \quad \textit{and} \quad q=\frac{(2-2p-2x)}{\alpha(1-2p)}. \tag{4.4}$$

Then

$$\begin{aligned} t &= 2\sqrt{-\frac{r}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2r}\sqrt{\frac{-3}{r}}\right)-\frac{2\pi}{3}\right)= \\ &= -2\sqrt{-\frac{r}{3}}\sin\left(\frac{\pi}{6}-\frac{1}{3}\arccos\left(\frac{3q}{2r}\sqrt{\frac{-3}{r}}\right)\right) \end{aligned} \tag{4.5}$$

is a real root in the interval $[0,1]$ because $r<0$ for any p and α satisfying condition (2.2). The left hand side of (4.1) is decreasing function, hence we change the direction of inequality and we obtain

$$\begin{aligned} P(x\geq\pi(t)) &= P\left(t<-2\sqrt{-\frac{r}{3}}\sin\left(\frac{\pi}{6}-\frac{1}{3}\arccos\left(\frac{3q}{2r}\sqrt{\frac{-3}{r}}\right)\right)\right)= \\ &= -2\sqrt{-\frac{r}{3}}\sin\left(\frac{\pi}{6}-\frac{1}{3}\arccos\left(\frac{3q}{2r}\sqrt{\frac{-3}{r}}\right)\right). \end{aligned} \tag{4.6}$$

Next, we calculate the probability that a peaceful split $(\frac{1}{2},\frac{1}{2})$ is accepted. The joint cumulative distribution function of random variables $\pi(t_1)$ and $\pi(t_2)$ is

$$P(1/2\geq\pi(t_1)\wedge 1/2\geq\pi(t_2))=P(1/2\geq\pi(t_1))P(1/2\geq\pi(t_2))=P^2(1/2\geq\pi(t))$$

because $\pi(t_1)$ and $\pi(t_2)$ are independent and identically distributed. Hence

$$\mathscr{P}(1/2,1/2)=P^2(1/2\geq\pi(t))=\phi^2(1/2).$$

Substituting $x = 1/2$ in (4.4) we obtain $q = \frac{1}{\alpha}$. From (4.6) we obtain

$$P(1/2 \geq \pi(t)) = 2\sqrt{-\frac{r}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2r}\sqrt{\frac{-3}{r}}\right) - \frac{2\pi}{3}\right)$$

where

$$r = \frac{-2-\alpha+4p+\alpha p}{\alpha(1-2p)} \quad \textit{and} \quad q = \frac{1}{\alpha}. \tag{4.7}$$

Hence,

$$\mathscr{P}(1/2, 1/2) = -\frac{4}{3}r\sin^2\left(\frac{\pi}{6} - \frac{1}{3}\arccos\left(\frac{3\sqrt{3}q\sqrt{-\frac{1}{r}}}{2r}\right)\right)$$

where r and q are given by (4.7). $\square$

Proof of Proposition 7

The proof is by construction. We calculate parameters of the optimal mediation programme and we show that both types have no incentive to deviate from recommendations of the mediator.

1. We note that s_l appears only in left-hand side of the low type IC constraint (2.79) and this left-hand side is increasing in s_l because $\pi_l(t_i) \leq \frac{1}{2}$ by lemma (2). Hence, $s_l = 1$ in the solution of the problem (2.76).

2. We note that by lemma (2) the coefficient of s_h in the left-hand side of the high type IR constraint (2.78) is positive. Moreover, the coefficient $\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8})$ of s_h in the right-hand side of the low type *interim* IC constraint is also positive. Hence, the low type *interim* IC constraint must be binding for some $t_i \leq 1/2$ in the solution of relaxed problem, otherwise we could increase s_h thus increasing the value of the objective function, without violating the high type IR constraint (2.78).

3. We note that the high type IR constraint must be binding for some $t_i \geq 1/2$ in the solution of relaxed problem, otherwise we could decrease b and make the low type IC constraint slack.

4. We would like to show that $s_h > s_m$ in the solution of relaxed problem, which implies that the low type IC constraint binds for $t_i = 1/2$. Further, we show that the IR constraint for the high type binds for $t_i = 1$.

   In the light of step one, the constraint (2.61) which binds for some $t_i = t_i^*$ becomes

   $$\frac{1}{2} - \pi_l(t_i^*) + s_m\pi_l(t_i^*) - 2b + s_m\frac{1}{2} + s_m\left(\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha t_i^*}{8})\right) =$$

   $$= s_h\left(\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha t_i^*}{8})\right).$$

   As $\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha t_i}{8}) > 0$ for any t_i, parameter

   $$s_h = s_m + \frac{\frac{1}{2} - \pi_l(t_i^*) + s_m\pi_l(t_i^*) + s_m\frac{1}{2} - 2b}{\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha t_i^*}{8})} \tag{4.8}$$

   is well defined. We would like to show that the second term in the right-hand side of (4.8) is positive. Clearly, $\frac{1}{2} - (1-p)(\frac{1}{2} - \frac{3\alpha t_i^*}{8}) > 0$. It remains to show that inequality

$$\frac{1}{2} - \pi_l(t_i^*) + s_m \pi_l(t_i^*) + s_m \frac{1}{2} - 2b > 0 \tag{4.9}$$

holds.

The IR constraint (2.60) binds for some $t_i = t_i^{**}$, therefore

$$b = s_m p \left(\frac{1}{2} - \frac{\alpha t_i^{**}}{8} \right) - s_h \left(\frac{1}{2} - \pi_h(t_i^{**}) \right).$$

We note that $\frac{1}{2} - \pi_h(t_i^{**}) > 0$ by lemma (2). Hence, $b \leq s_m p \frac{1}{2}$ for any feasible $s_h \geq 0$.

Therefore, the left-hand side of (4.9) satisfies

$$\frac{1}{2} - \pi_l(t_i^*) + s_m \pi_l(t_i^*) + s_m \frac{1}{2} - 2b > \frac{1}{2} + (s_m - 1)\pi_l(t_i^*) + s_m \frac{1}{2} - s_m p. \tag{4.10}$$

We note that inequality $s_m - 1 \leq 0$ holds for the probability s_m. Hence, the right-hand side of (4.10) is minimal when the value of $\pi_l(t_i^*)$ is maximal, that is, for $t_i = 1/2$. Hence,

$$\frac{1}{2} + (s_m - 1)\pi_l(t_i^*) + s_m \frac{1}{2} - s_m p \geq \frac{1}{2} + (s_m - 1)\frac{p}{2}(1 - \frac{\alpha}{8}) + s_m(\frac{1}{2} - p) =$$

$$= \frac{1}{2} - \frac{p}{2} + \frac{\alpha p}{16} + s_m(\frac{p}{2}(1 - \frac{\alpha}{8}) + \frac{1}{2} - p) = \frac{1}{2} - \frac{p}{2} + \frac{\alpha p}{16} + s_m(\frac{1}{2} - \frac{p}{2} - \frac{\alpha p}{16}) =$$
$$= g(s_m).$$

For parameter values for which $\frac{1}{2} - \frac{p}{2} - \frac{\alpha p}{16} > 0$ the value of $g(s_m)$ is minimal for $s_m = 0$. Then $g(0) = \frac{1}{2} - \frac{p}{2} + \frac{\alpha p}{16} > 0$ for any parameters values. For parameter values for which $\frac{1}{2} - \frac{p}{2} - \frac{\alpha p}{16} < 0$ the value of $g(s_m)$ is minimal for $s_m = 1$. Then the minimal value of $g(s_m)$ is $g(1) = 1 - p > 0$. Hence, for any parameters values $g(s_m) > 0$ and inequality (4.9) is satisfied.

Therefore, in the light of equation (4.8) we conclude that $s_h > s_m$ in the solution of relaxed problem. In the light of step 1 we rewrite the *interim* IC constraint for the low type (2.61) as

$$\frac{1}{2} + s_m - s_m(1 - p)\frac{1}{2} \geq$$

$$\geq 2b+(1-s_m)\pi_l(t_i)+s_h\frac{1}{2}-s_h(1-p)\frac{1}{2}+(s_h-s_m)(1-p)\frac{3\alpha t_i}{8}.$$

We notice that the right hand side of this inequality is maximal for $t_i = 1/2$ because $1-s_m \geq 0$, $s_h - s_m > 0$, and by lemma (2) $\pi_l(t_i)$ is maximal for $t_i = 1/2$. Therefore, the *interim* IC constraint for the low type binds for $t_i = 1/2$.

It is easy to check that for $s_h > s_m$ and $t_i \in [1/2, 1]$ the value of expression $s_m p\frac{\alpha t_i}{8} - s_h \pi_h(t_i)$ in the left hand side of IR constraint (2.78) is minimal for $t_i = 1$. Hence, the IR constraint for the high type binds for $t_i = 1$. Therefore, the IR constraint

$$s_h\left(\frac{1}{2}-\pi_h(1)\right)+b-s_m p\left(\frac{1}{2}-\frac{\alpha}{8}\right)\geq 0$$

and the IC constraint

$$\frac{1}{2}-\pi_l(1/2)+s_m-s_m(1-p)\left(\frac{1}{2}-\frac{3\alpha}{16}\right)\geq$$

$$\geq 2b-s_m\pi_l(1/2)+s_h\left(\frac{1}{2}-(1-p)\left(\frac{1}{2}-\frac{3\alpha}{16}\right)\right)$$

bind.

5. We rewrite the constraints of the relaxed problem by substituting

$$b = s_m p(\frac{1}{2}-\frac{\alpha}{8}) - s_h\left(\frac{1}{2}-p(\frac{1}{2}-\frac{3\alpha}{8})\right)$$

in the IC constraint for $t_i = 1/2$

$$\frac{1}{2}-\frac{1}{2}p(1-\frac{\alpha}{8})+s_m-s_m(1-p)(\frac{1}{2}-\frac{3\alpha}{16})=$$

$$=2b-s_m\frac{1}{2}p(1-\frac{\alpha}{8})+s_h(\frac{1}{2}-(1-p)(\frac{1}{2}-\frac{3\alpha}{16}))$$

and we obtain

$$s_m = -\frac{8-8p+\alpha p}{8+3\alpha} - s_h\frac{16-3\alpha-24p+15\alpha p}{8+3\alpha} \tag{4.11}$$

In the light of step 1 we simplify the objective function (2.58) and maximize

$$max_{s_h,s_m}\left\{2s_m+s_h\right\}. \tag{4.12}$$

Substituting s_m by the RHS of (4.11) we maximize expression

$$W = const + s_h \frac{48p - 24 + 9\alpha - 30\alpha p}{8 + 3\alpha}.$$

We note that coefficient of s_h is positive for any values of parameters α and p that satisfy condition (2.2). Hence, the maximization of W requires maximization of s_h. The value of s_h is bounded by probability constraints (2.77). The *ex ante* probability of peace W in the Bayesian equilibrium of the induced game is maximal for $s_h = 1$. It is a feasible value for s_h for some range of parameters α and p.

We distinguish two cases for these parameter values. Clearly, (2.2) implies $\alpha \in (0, 1/2)$. For $\frac{1}{6}(19 - \sqrt{265}) < \alpha \le 1/2$ or

$$0 < \alpha \le \frac{1}{6}(19 - \sqrt{265}) \;\wedge\; \frac{24 - 3\alpha}{32 - 16\alpha} < p \le 1,$$

where $\frac{1}{6}(19 - \sqrt{265}) \approx 0.45353$, the value of 1 is feasible for s_h. For this range of parameters α and p we compute s_m from equation (4.11) and we obtain

$$0 < s_m = \frac{32p - 24 + 3\alpha - 16\alpha p}{8 + 3\alpha} < 1. \tag{4.13}$$

For $0 < \alpha \le \frac{1}{6}(19 - \sqrt{265}) \;\wedge\; 1/2 < p < \frac{24-3\alpha}{32-16\alpha}$ the maximal feasible value of s_h satisfying all constraints of the relaxed problem is the value corresponding to $s_m = 0$. This is a feasible value of s_m. However, the corresponding value of s_h and, hence, the value of W is lower for this range of parameter values.

We show in subsequence that for parameter values satisfying condition (4.13) the *ex ante* probability of peace in the Bayesian equilibrium of the induced game is lower than the *ex ante* probability of peace in the non-mediated game.

Let the solution of the initial problem (2.51) be $q_h = 1$,

$$q_m + p_m = \frac{32p - 24 + 3\alpha - 16\alpha p}{8 + 3\alpha}$$

and $q_l + 2p_l = 1$. In this case diads (h, h) and (l, l) do not fight in the Bayesian equilibrium of the induced game. We can choose the value of x in such a way that this solution does not violate the high-type IC constraint and the low type IR constraint. However, the *ex ante* probability of peace in the best separating equilibrium of the mediation game is for these values of parameters p and α the same as the *ex ante* probability of peace in the best separating equilibrium of the mediation game induced by mediation programme (2.50).

□

Chapter 5

Appendix B

Proof of Proposition 10

By a well known result of Hardy, Littlewood and Polya (1952) published in [7], for any agent i, the rank of the optimal weight's product $\lambda_i(t_i)f_i(t_i)$ in the increasing order of weights' products $\{\lambda_i(t_i)f_i(t_i)\}_{t_i \in T_i}$ is the same as the rank of the interim expected utility $U^i(\phi, t_i)$ in the increasing order of the interim expected utilities $\{U^i(\phi, t_i)\}_{t_i \in T_i}$. Since Γ is Bayesian incentive compatible, by Corollary 1.1 for any agent i and any three types $t_i'' \preceq_i t_i \preceq_i t_i'$, the interim expected utility $U^i(\phi, t_i)$ satisfies $U^i(\phi|t_i'') \leq U^i(\phi|t_i) \leq U^i(\phi|t_i')$, which establishes the result.

$\square$